

# Change of variable in spaces of mixed smoothness and numerical integration of multivariate functions on the unit cube

Van Kien Nguyen

Mario Ullrich

Tino Ullrich

November 9, 2015

## Abstract

In a recent article by two of the present authors it turned out that Frolov’s cubature formulae are optimal and universal for various settings (Besov-Triebel-Lizorkin spaces) of functions with dominating mixed smoothness. Those cubature formulae go well together with functions supported inside the unit cube  $[0, 1]^d$ . The question for the optimal numerical integration of multivariate functions with non-trivial boundary data, in particular non-periodic functions, arises. In this paper we give a general result that the asymptotic rate of the minimal worst-case integration error is not affected by boundary conditions in the above mentioned spaces. In fact, we propose two tailored modifications of Frolov’s cubature formulae suitable for functions supported on the cube (not in the cube) which provide the same minimal worst-case error up to a constant. This constant involves the norms of a “change of variable” and a “pointwise multiplication” mapping, respectively, between the function spaces of interest. In fact, we complement, extend and improve classical results by Bykovskii, Dubinin and Temlyakov on the boundedness of change of variable mappings in Besov-Sobolev spaces of mixed smoothness. Our proof technique relies on a new characterization via integral means of mixed differences and maximal function techniques, general enough to treat Besov and Triebel-Lizorkin spaces at once. The second modification, which only tackles the case of periodic functions, is based on a pointwise multiplication and is therefore most likely more suitable for applications than the (traditional) “change of variable” approach. These new theoretical insights are expected to be useful for the design of new (and robust) cubature rules for multivariate functions on the cube.

## 1 Introduction

The seminal papers by Korobov [16], Hlawka [15] and Bakhvalov [2] started the research for the efficient computation of multivariate integrals  $\int_{\Omega} f(x) dx$  via cubature formulae of type

$$Q_n(f) := \sum_{i=1}^n \lambda_i f(x^i), \quad (1.1)$$

where  $X_n := \{x^1, \dots, x^n\} \subset \Omega \subset \mathbb{R}^d$  denotes the set of given integration nodes and  $(\lambda_1, \dots, \lambda_n)$  denotes the vector of integration weights. This field developed into several directions and attracted a lot of interest in the past 50 years until present, see Temlyakov [28, 29, 30, 32], Dubinin [7, 8], Skriganov [26], Triebel [36], Hinrichs et al. [13, 14], Novak and Woźniakowski [21], Novak, Krieg [17], Dick and Pillichshammer [6], Dũng, Ullrich [9], and Markhasin [19] to mention just a few. However, several fundamental problems, for instance the construction of optimal (and easy to generate) point sets  $X_n$  in (1.1) is still subject of intense research. Frolov’s construction [11] dates back to the 1970s and gives a relatively easy construction of

such point sets which perform optimal for several function classes of mixed smoothness. The recent results in [38] will be the starting point for the investigations in this paper.

The concept of *dominating mixed smoothness* does not only connect discrepancy theory and optimal numerical integration. Function classes built upon this type of multivariate regularity also play an important role in many real-world problems. Several applications are modeled in Sobolev spaces with dominating mixed smoothness  $H_{\text{mix}}^s$ , see for instance Yserentant's book [43] for regularity properties of solutions of the electronic Schrödinger equation. However, the Hilbert space model is not sufficient when it comes to numerical integration issues. We observed in [38, 14] that it seems to be more appropriate to define smoothness function classes on  $L_p$  instead of  $L_2$ . The case  $p = 1$  turns out to be particularly important. To see this let us consider multi-variate *kink functions* which often occur in mathematical finance, e.g., the pricing of a European call option, whose pay-off function possesses a kink at the strike price [12]. In general, one can not expect Sobolev regularity higher than  $s = 3/2$ . However, when considering Besov regularity we can achieve smoothness  $s = 2$ . In a sense, one sacrifices integrability for gaining regularity. Looking at the error bounds in [38, 14] one observes that certain cubature rules benefit from higher Besov regularity while the integrability  $p$  does not enter the picture.

In the recent paper [38] two of the present authors studied the performance of the classical Frolov cubature formulae [11, 37] in several settings of classes with dominating mixed smoothness. It turned out that in Besov  $(\mathring{\mathbf{B}}_{p,\theta}^s)$  and Triebel-Lizorkin  $(\mathbf{F}_{p,\theta}^s)$  spaces with dominating mixed smoothness  $\mathbf{F}_d := \mathring{\mathbf{A}}_{p,\theta}^s$ , where  $\mathbf{A} \in \{\mathbf{B}, \mathbf{F}\}$ , of functions supported strictly inside the unit cube  $\Omega := [0, 1]^d$ , the above mentioned cubature formulae provide the optimal rate of convergence among all cubature formulae of type (1.1) with respect to the minimal worst-case integration error given by

$$\text{Int}_n(\mathbf{F}_d) := \inf_{Q_n} e(Q_n, \mathbf{F}_d), \quad n \in \mathbb{N}, \quad (1.2)$$

where the infimum is taken over all cubature formulae of the form (1.1) and

$$e(Q_n, \mathbf{F}_d) := \sup_{\|f\|_{\mathbf{F}_d} \leq 1} |I(f) - Q_n(f)|, \quad n \in \mathbb{N},$$

for a class  $\mathbf{F}_d \subset C(\mathbb{R}^d)$  with  $I(f) := \int_{[0,1]^d} f(x) dx$ . However, Frolov's method does not seem to be suitable for functions with non-trivial boundary data, like multivariate periodic functions  $\mathbf{F}_d := \mathbf{A}_{p,\theta}^s(\mathbb{T}^d)$  or non-periodic functions on the cube  $\mathbf{F}_d := \mathbf{A}_{p,\theta}^s([0, 1]^d)$ , cf. Definitions 3.1, 3.2 below. The case of periodic Besov spaces of dominating mixed smoothness has been studied in [7, 8, 9, 14]. What concerns optimal cubature in non-periodic function classes, which are actually most relevant for applications, there are only partial results known, see for instance [36, 9] for the correct lower bounds of  $\text{Int}_n(\mathbf{B}_{p,\theta}^s)$ .

A classical approach towards upper bounds has been proposed by Bykovskii [3] by first performing a change of variable to obtain

$$\int_{[0,1]^d} f(x) dx = \int_{[0,1]^d} |\det \psi'(x)| f(\psi(x)) dx \quad (1.3)$$

for some differentiable  $\psi : [0, 1]^d \mapsto [0, 1]^d$ , and use afterwards a cubature formula (1.1) for the right-hand integrand in (1.3). The main observation is the fact that this approach results in a modified cubature formula

$$Q_n^\psi(f) := Q_n(|\det \psi'| \cdot f \circ \psi) = \sum_{i=1}^n \lambda_i |\det \psi'(x^i)| f(\psi(x^i)). \quad (1.4)$$

Arranging the kernel  $\psi$  in a way such that the function  $|\det \psi'(x)|f(\psi(x))$  preserves mixed smoothness properties and is supported in  $[0, 1]^d$  one can now use the above mentioned (optimal) methods to approximate the integral. Thus, the remaining step in analyzing the method (1.4) is to prove the preservation of mixed smoothness properties under change of variable, or in other words, to establish the boundedness of mappings

$$\begin{aligned} T_d^\psi : \mathbf{A}_{p,\theta}^s &\rightarrow \mathbf{A}_{p,\theta}^s \\ f &\mapsto |\det \psi'(x)|f(\psi(x)), \end{aligned} \quad (1.5)$$

for suitably chosen  $\psi(x)$ . A straight-forward choice is certainly given by assuming a tensor product structure. That is, e.g., given univariate functions  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  which are integrated  $C^k(\mathbb{R})$  bump functions  $\varphi$  supported in  $[0, 1]$  and, with a slight abuse of notation, define  $\psi(x_1, \dots, x_d) := \psi(x_1) \cdot \dots \cdot \psi(x_d)$ . A natural and simple choice is given by the family of polynomials

$$\psi_k(t) = \begin{cases} \int_0^t \xi^k(1-\xi)^k d\xi / \int_0^1 \xi^k(1-\xi)^k d\xi & : t \in [0, 1], \\ 1 & : t > 1, \\ 0 & : t < 0, \end{cases} \quad (1.6)$$

$k \in \mathbb{N}$  and we define

$$\psi_k(x) = \psi_k(x_1) \cdot \dots \cdot \psi_k(x_d). \quad (1.7)$$

It has been proved by Bykovskii [3] that  $T_d^{\psi_k}$  is bounded in the Sobolev space with dominating mixed smoothness  $\mathbf{W}_2^s$  for  $s \in \mathbb{N}$  if  $m \geq 2[s] + 1$  (note that this is the same space as  $H_{\text{mix}}^s$  mentioned above). This result has been extended by Temlyakov, see [31, Thm. IV.4.1], to Sobolev spaces  $\mathbf{W}_p^s$ ,  $s \in \mathbb{N}$ ,  $1 < p < \infty$ , under the condition

$$k \geq \left\lfloor \frac{sp}{p-1} \right\rfloor + 1. \quad (1.8)$$

Under the same condition, Temlyakov [31, Lem. IV.4.9] showed the boundedness of  $T_d^{\psi_k} : \mathbf{B}_{p,\infty}^s \rightarrow \mathbf{B}_{p,\infty}^s$  in the Hölder-Nikol'skij spaces if  $1 < p \leq \infty$  and  $s > 1$ . Later this was extended by Dubinin [7, 8] to the Besov spaces  $\mathbf{B}_{p,\theta}^s$  if  $1 \leq p, \theta \leq \infty$  and  $s > 1/p$ , and by Temlyakov [32] to the spaces  $\mathbf{W}_p^s$ ,  $s \in \mathbb{N}$ ,  $1 \leq p \leq \infty$ . Dubinin and Temlyakov analyzed the change of variable with a  $C^\infty$ -kernel instead of (1.6), namely

$$\psi(t) = \begin{cases} \int_0^t e^{-\frac{1}{\xi(1-\xi)}} d\xi / \int_0^1 e^{-\frac{1}{\xi(1-\xi)}} d\xi & : t \in [0, 1], \\ 1 & : t > 1, \\ 0 & : t < 0. \end{cases} \quad (1.9)$$

With this particular choice both authors were able to incorporate even the important case  $p = 1$  (see the discussion above). Note, that the boundedness is still open for  $\mathbf{F}_{1,\theta}^s$ , see Remark 4.3, (iii) below. However, the kernel in (1.9) has some obvious disadvantages. It involves terms which get very small and others that get very large at the same time. Therefore, motivated from numerical robustness issues one is interested in as simple as possible change of variable kernels with the least (smoothness) requirements. The functions in (1.6) are polynomials and therefore good for our purpose. However, from a *tractability* point of view, i.e., if one is interested in the dependence of the involved constants on the dimension, every method of this type will be most likely intractable. Some results on this issue in the Hilbert space setting can be found in Kuo, Sloan and Woźniakowski [18], see also Nuyens and Cools [22] for numerical experiments.

Our main intention is to extend the mentioned boundedness results on the operators  $T_d^{\psi_k}$  to spaces  $\mathbf{A}_{p,\theta}^s$ , in particular Triebel-Lizorkin spaces of dominating mixed smoothness  $\mathbf{F}_{p,\theta}^s$ , for the largest possible range of parameters.

## 1.1 Contribution and main results

The main goal of the present paper is to construct cubature formulae that

- are simple, (to some extent) stable and universal (in the sense that the algorithm does not depend on the specific parameters of the spaces)
- and that perform optimal in the sense of the minimal worst-case error (1.2) with respect to non-periodic function spaces  $\mathbf{A}_{p,\theta}^s$  on the unit  $d$ -cube.

The ingredient (1.6) is almost what we want to put into (1.4), provided we can get rid of the condition (1.8). This is possible as we will show in the theorems below. We are able to state the result for spaces  $\mathbf{A}_{p,\theta}^s$  on the unit  $d$ -cube, where the smoothness (or the parameter  $k$  in (1.6)) only depends on the smoothness  $s$  of the space (and not on the integrability  $p$ ). Note, that the result below is not just a generalization of the results from [3, 8, 31, 32]. It essentially improves on the stated classical results for  $\mathbf{W}_p^s$  and  $\mathbf{B}_{p,\theta}^s$  what concerns the choice of the change of variable kernel, which is an essential ingredient for the resulting modified cubature formula.

An important special case of our main result in Theorem 4.1 below gives the desired result for the Besov spaces  $\mathbf{B}_{p,\theta}^s$  of dominating mixed smoothness.

**Theorem 1.1.** Let  $\Omega = [0, 1]^d$ ,  $1 \leq p \leq \infty$ ,  $0 < \theta \leq \infty$  and  $s > 0$ . Let further  $k > \lfloor s \rfloor + 2$  ( $k > \lfloor s \rfloor + 3$  if  $p = 1$ ) and  $\psi_k$  as in (1.7). Then,

- (i)  $T_d^{\psi_k} : \mathbf{B}_{p,\theta}^s \rightarrow \mathbf{B}_{p,\theta}^s$  is a bounded mapping, and
- (ii) provided that  $\mathbf{B}_{p,\theta}^s \subset C(\mathbb{R}^d)$ , a corresponding modified cubature formula (1.4) on  $\mathbf{B}_{p,\theta}^s$  does not perform asymptotically worse than (1.1) performs on  $\mathring{\mathbf{B}}_{p,\theta}^s$ , i.e.,

$$e(Q_n^{\psi_k}, \mathbf{B}_{p,\theta}^s) \lesssim e(Q_n, \mathring{\mathbf{B}}_{p,\theta}^s), \quad n \in \mathbb{N}.$$

The more restrictive condition in the case  $p = 1$  is actually not necessary, see Remark 4.7, but a proof of this statement would elongate the paper by some pages, so we leave it out.

An analogous results for the Triebel-Lizorkin spaces of dominating mixed smoothness  $\mathbf{F}_{p,\theta}^s$  follows from Theorem 4.2 below. Note that we have to exclude the corner points  $p = 1, \infty$ .

**Theorem 1.2.** Let  $\Omega = [0, 1]^d$ ,  $1 < p < \infty$ ,  $1 < \theta \leq \infty$  and  $s > 0$ . Let further  $k > \lfloor s \rfloor + 2$  and  $\psi_k$  as in (1.7). Then,

- (i)  $T_d^{\psi_k} : \mathbf{F}_{p,\theta}^s \rightarrow \mathbf{F}_{p,\theta}^s$  is a bounded mapping, and
- (ii) provided that  $\mathbf{F}_{p,\theta}^s \subset C(\mathbb{R}^d)$ , a corresponding modified cubature formula (1.4) on  $\mathbf{F}_{p,\theta}^s$  does not perform asymptotically worse than (1.1) performs on  $\mathring{\mathbf{F}}_{p,\theta}^s$ , i.e.,

$$e(Q_n^{\psi_k}, \mathbf{F}_{p,\theta}^s) \lesssim e(Q_n, \mathring{\mathbf{F}}_{p,\theta}^s), \quad n \in \mathbb{N}.$$

These results are useful in the sense that they “transfer” the optimal rate of the minimal worst-case error from the known situation  $\mathring{\mathbf{A}}_{p,\theta}^s$  to the more difficult situation  $\mathbf{A}_{p,\theta}^s$ . As a direct consequence we obtain the equivalence

$$\text{Int}_n(\mathring{\mathbf{A}}_{p,\theta}^s) \asymp \text{Int}_n(\mathbf{A}_{p,\theta}^s(\mathbb{T}^d)) \asymp \text{Int}_n(\mathbf{A}_{p,\theta}^s([0, 1]^d)), \quad n \in \mathbb{N},$$

in the range of parameters given in Theorems 1.1, 1.2 whenever the classes  $\mathbf{A}_{p,\theta}^s$  consist of continuous functions. Let us mention the important special case

$$\text{Int}_n(\dot{\mathbf{W}}_p^s) \asymp \text{Int}_n(\mathbf{W}_p^s(\mathbb{T}^d)) \asymp \text{Int}_n(\mathbf{W}_p^s([0,1]^d)), \quad n \in \mathbb{N},$$

for  $1 < p < \infty$  and  $s > 1/p$  taking into account that  $\mathbf{F}_{p,2}^s = \mathbf{W}_p^s$ . Note, that this includes the case of small smoothness, which is present if  $2 < p < \infty$  and  $1/p < s \leq 1/2$ , see [30], [38].

In case, one is only interested in the numerical integration of periodic functions, we present an even simpler modification that satisfies the second point above but uses only a pointwise multiplication with a certain function. To our best knowledge this method has not been used before. In contrast to the considerations above we will now restrict to function spaces on the  $d$ -torus  $\mathbb{T}^d$ , see Subsection 3.2 below, i.e., we want to numerically integrate functions which are 1-periodic in every component. There are many results for  $\mathbf{W}_p^s(\mathbb{T}^d)$  and  $\mathbf{B}_{p,\infty}^s(\mathbb{T}^d)$  in this direction, see for instance [32] and the references therein. The recent contribution [14] especially deals with periodic Besov spaces of dominating mixed smoothness and analyzes cubature formulae on digital nets. However, the picture is still not complete. Except for the Fibonacci cubature rule [30, 9] in  $d = 2$  the discussed methods in [36, 14] are either not optimal or the optimal methods only work for specific (restricted) sets of parameters, like for instance small smoothness  $s < 2$  in [14]. In other words, the cubature formula is optimal for low smoothness but is not able to benefit from higher regularity. One goal of this paper is to get rid of such conditions. To gain this, we pursue a similar goal as above, namely the reduction of the problem to the  $\dot{\mathbf{A}}_{p,\theta}^s$  setting via a tailored transformation. To this end we use a sufficiently smooth function  $\psi : \mathbb{R}^d \rightarrow [0, \infty)$  being compactly supported ( $\text{supp}(\psi) \subset \Omega$ ) such that

$$\sum_{\ell \in \mathbb{Z}^d} \psi(x + \ell) = 1 \quad , \quad x \in \mathbb{R}^d. \quad (1.10)$$

Clearly, we need  $\Omega \supsetneq [0,1]^d$ . Then for any 1-periodic integrable function in each component it holds

$$\begin{aligned} \int_{\mathbb{R}^d} \psi(x) f(x) dx &= \sum_{\ell \in \mathbb{Z}^d} \int_{[0,1]^d} \psi(x + \ell) f(x + \ell) dx \\ &= \int_{[0,1]^d} f(x) \sum_{\ell \in \mathbb{Z}^d} \psi(x + \ell) dx = \int_{[0,1]^d} f(x) dx. \end{aligned}$$

Starting with a cubature formula  $Q_n$  of type (1.1) for functions with support in  $\Omega$  we modify as follows

$$\tilde{Q}_n^\psi := Q_n(\psi f) = \sum_{i=1}^n \psi(x^i) \lambda_i f(\{x^i\}), \quad (1.11)$$

where  $\{x\} = x - [x]$  has to be understood component-wise. In contrast to (1.5) the remaining step for analyzing (1.11) on 1-periodic function spaces is to establish the boundedness of pointwise multiplier mappings

$$\begin{aligned} \tilde{T}_d^\psi : \mathbf{A}_{p,\theta}^s(\mathbb{T}^d) &\rightarrow \mathbf{A}_{p,\theta}^s(\mathbb{R}^d) \\ f(x) &\mapsto \psi(x) f(x). \end{aligned} \quad (1.12)$$

There is a rich theory concerning pointwise multiplication for isotropic spaces  $B_{p,\theta}^s(\mathbb{R}^d)$ ,  $F_{p,\theta}^s(\mathbb{R}^d)$ , see Runst, Sickel [23]. To our best knowledge the field is rather open for spaces of dominating mixed smoothness. In contrast to the theorems above on change of variable we are able to

include the quasi-Banach case ( $\min\{p, \theta\} < 1$ ) within this framework. This is why the natural lower smoothness restrictions enter, namely  $s > \sigma_p$  and  $s > \sigma_{p,\theta}$ , respectively, where

$$\sigma_p := \max\left\{\frac{1}{p} - 1, 0\right\} \quad \text{and} \quad \sigma_{p,\theta} := \max\left\{\frac{1}{p} - 1, \frac{1}{\theta} - 1, 0\right\}.$$

For the  $\mathbf{B}$ -spaces the result looks as follows.

**Theorem 1.3.** Let  $0 < p, \theta \leq \infty$ ,  $s > \sigma_p$  and  $\psi \in C_0^k(\Omega)$  with  $k \geq \lfloor s \rfloor + 1$ . Then,

- (i)  $\tilde{T}_d^\psi : \mathbf{B}_{p,\theta}^s(\mathbb{T}^d) \rightarrow \mathbf{B}_{p,\theta}^s$  is a bounded mapping, and
- (ii) provided that  $\mathbf{B}_{p,\theta}^s \subset C(\mathbb{R}^d)$  and  $\psi$  satisfies (1.10), a corresponding modified cubature formula (1.11) on  $\mathbf{B}_{p,\theta}^s(\mathbb{T}^d)$  does not perform asymptotically worse than (1.1) performs on  $\mathring{\mathbf{B}}_{p,\theta}^s(\Omega)$ , i.e.,

$$e(\tilde{Q}_n^\psi, \mathbf{B}_{p,\theta}^s(\mathbb{T}^d)) \lesssim e(Q_n, \mathring{\mathbf{B}}_{p,\theta}^s(\Omega)), \quad n \in \mathbb{N}.$$

The result for the  $\mathbf{F}$ -spaces is similar, but we have to exclude the case  $p = \infty$ .

**Theorem 1.4.** Let  $0 < p < \infty$ ,  $0 < \theta \leq \infty$ ,  $s > \sigma_{p,\theta}$  and  $\psi \in C_0^k(\Omega)$  with  $k \geq \lfloor s \rfloor + 1$ . Then,

- (i)  $\tilde{T}_d^\psi : \mathbf{F}_{p,\theta}^s(\mathbb{T}^d) \rightarrow \mathbf{F}_{p,\theta}^s$  is a bounded mapping, and
- (ii) provided that  $\mathbf{F}_{p,\theta}^s \subset C(\mathbb{R}^d)$  and  $\psi$  satisfies (1.10), a corresponding modified cubature formula (1.11) on  $\mathbf{F}_{p,\theta}^s(\mathbb{T}^d)$  does not perform asymptotically worse than (1.1) performs on  $\mathring{\mathbf{F}}_{p,\theta}^s(\Omega)$ , i.e.,

$$e(\tilde{Q}_n^\psi, \mathbf{F}_{p,\theta}^s(\mathbb{T}^d)) \lesssim e(Q_n, \mathring{\mathbf{F}}_{p,\theta}^s(\Omega)), \quad n \in \mathbb{N}.$$

Both theorems immediately imply the two-sided relation

$$\text{Int}_n(\mathring{\mathbf{A}}_{p,\theta}^s) \asymp \text{Int}_n(\mathbf{A}_{p,\theta}^s(\mathbb{T}^d)), \quad n \in \mathbb{N}, \quad (1.13)$$

in the above range of parameters if additionally  $\mathbf{A}_{p,\theta}^s \hookrightarrow C(\mathbb{R}^d)$  holds true.

## 1.2 Organization of the paper

The paper is organized as follows. In Section 2 we collect the required tools from Fourier analysis, especially maximal inequalities. Section 3 is devoted to the definition and properties of the function spaces under consideration. There we characterize Besov-Triebel-Lizorkin spaces by mixed iterated differences which turn out to be the crucial tool for proving our main results. Sections 4 and 5 are the heart of the paper in which we prove the change of variable and pointwise multiplication results already discussed in the introduction section. For the convenience of the reader we also discuss the one-dimensional versions of the results in Sections 4 and 5 in detail.

## 1.3 Notation

As usual  $\mathbb{N}$  denotes the natural numbers,  $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ ,  $\mathbb{Z}$  denotes the integers,  $\mathbb{R}$  the real numbers, and  $\mathbb{C}$  the complex numbers. By  $\mathbb{T} := \mathbb{R}/\mathbb{Z}$  we denote the torus represented with the interval  $[0, 1]$  in  $\mathbb{R}$ . The letter  $d$  is always reserved for the underlying dimension in  $\mathbb{R}^d, \mathbb{Z}^d$  etc and by  $[d]$  we mean  $[d] = \{1, \dots, d\}$ . We denote with  $\langle x, y \rangle$  or  $x \cdot y$  the usual Euclidean inner product in  $\mathbb{R}^d$  or  $\mathbb{C}^d$ . For  $0 < p \leq \infty$  and  $x \in \mathbb{R}^d$  we denote  $|x|_p = (\sum_{i=1}^d |x_i|^p)^{1/p}$  with the

usual modification in the case  $p = \infty$ . By  $(x_1, \dots, x_d) > 0$  we mean that each coordinate is positive. As usual we decompose  $x \in \mathbb{R}$  in  $x = \lfloor x \rfloor + \{x\}$  where  $0 \leq \{x\} < 1$  (resp.  $\{x\}$  for  $x \in \mathbb{R}^d$  is meant component-wise).

If  $X$  and  $Y$  are two (quasi-)normed spaces, the (quasi-)norm of an element  $x$  in  $X$  will be denoted by  $\|x\|_X$ . The symbol  $X \hookrightarrow Y$  indicates that the identity operator is continuous. For two sequences  $a_n$  and  $b_n$  we will write  $a_n \lesssim b_n$  if there exists a constant  $c > 0$  such that  $a_n \leq c b_n$  for all  $n$ . We will write  $a_n \asymp b_n$  if  $a_n \lesssim b_n$  and  $b_n \lesssim a_n$ .

## 2 Tools from Fourier analysis

In this section we will collect the required tools from Fourier analysis.

### 2.1 Preliminaries

Let  $L_p = L_p(\mathbb{R}^d)$ ,  $0 < p \leq \infty$ , be the space of all functions  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  such that

$$\|f\|_p := \left( \int_{\mathbb{R}^d} |f(x)|^p dx \right)^{1/p} < \infty$$

with the usual modification if  $p = \infty$ . We will also need  $L_p$ -spaces on compact domains  $\Omega \subset \mathbb{R}^d$  instead of  $\mathbb{R}^d$ . We write  $\|f\|_{L_p(\Omega)}$  for the corresponding (restricted)  $L_p$ -norm.

For  $k \in \mathbb{N}_0$  we denote by  $C_0^k(\mathbb{R}^d)$  the collection of all compactly supported functions  $\varphi$  on  $\mathbb{R}^d$  which have  $D^\alpha \varphi$  uniformly continuous on  $\mathbb{R}^d$  for  $|\alpha|_\infty \leq k$ . Additionally, we define the spaces infinitely differentiable functions  $C^\infty(\mathbb{R}^d)$  and infinitely differentiable functions with compact support  $C_0^\infty(\mathbb{R}^d)$  as well as the *Schwartz space*  $\mathcal{S} = \mathcal{S}(\mathbb{R}^d)$  of all rapidly decaying infinitely differentiable functions on  $\mathbb{R}^d$ , i.e.,

$$\mathcal{S} := \{ \varphi \in C^\infty(\mathbb{R}^d) : \|\varphi\|_{k,\ell} < \infty \text{ for all } k, \ell \in \mathbb{N} \},$$

and

$$\|\varphi\|_{k,\ell} := \left\| (1 + |\cdot|)^k \sum_{|\alpha|_1 \leq \ell} |D^\alpha \varphi(\cdot)| \right\|_\infty, \quad \alpha \in \mathbb{N}_0^d.$$

The space  $\mathcal{S}'(\mathbb{R}^d)$ , the topological dual of  $\mathcal{S}(\mathbb{R}^d)$ , is also referred to as the set of tempered distributions on  $\mathbb{R}^d$ . Indeed, a linear mapping  $f : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathbb{C}$  belongs to  $\mathcal{S}'(\mathbb{R}^d)$  if and only if there exist numbers  $k, \ell \in \mathbb{N}$  and a constant  $c = c_f$  such that

$$|f(\varphi)| \leq c_f \|\varphi\|_{k,\ell}$$

for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . The space  $\mathcal{S}'(\mathbb{R}^d)$  is equipped with the weak\*-topology.

For  $f \in L_1(\mathbb{R}^d)$  we define the Fourier transform

$$\mathcal{F}f(\xi) = \int_{\mathbb{R}^d} f(y) e^{-2\pi i \langle \xi, y \rangle} dy, \quad \xi \in \mathbb{R}^d,$$

and the corresponding inverse Fourier transform  $\mathcal{F}^{-1}f(\xi) = \mathcal{F}f(-\xi)$ . As usual, the Fourier transform can be extended to  $\mathcal{S}'(\mathbb{R}^d)$  by  $(\mathcal{F}f)(\varphi) := f(\mathcal{F}\varphi)$ , where  $f \in \mathcal{S}'(\mathbb{R}^d)$  and  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . The mapping  $\mathcal{F} : \mathcal{S}'(\mathbb{R}^d) \rightarrow \mathcal{S}'(\mathbb{R}^d)$  is a bijection.

The convolution  $\varphi * \psi$  of two square-integrable functions  $\varphi, \psi$  is defined via the integral

$$(\varphi * \psi)(x) = \int_{\mathbb{R}^d} \varphi(x - y) \psi(y) dy.$$

If  $\varphi, \psi \in \mathcal{S}(\mathbb{R}^d)$  then  $\varphi * \psi$  still belongs to  $\mathcal{S}(\mathbb{R}^d)$ . In fact, we have  $\varphi * \psi \in \mathcal{S}(\mathbb{R}^d)$  even if  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  and  $f \in L_1(\mathbb{R}^d)$ . The convolution can be extended to  $\mathcal{S}(\mathbb{R}^d) \times \mathcal{S}'(\mathbb{R}^d)$  via  $(\varphi * f)(x) = f(\varphi(x - \cdot))$ . It makes sense pointwise and is a  $C^\infty$ -function in  $\mathbb{R}^d$ .

## 2.2 Maximal inequalities

Let us provide here the maximal inequalities for the Hardy-Littlewood and Peetre maximal functions, respectively, which are essential for the proofs of Theorems 4.1 and 4.2. For further details we refer to [42, 1.2, 1.3] or [24, Chapt. 2].

For a locally integrable function  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  we denote by  $Mf(x)$  the Hardy-Littlewood maximal function defined by

$$(Mf)(x) = \sup_{x \in Q} \frac{1}{|Q|} \int_Q |f(y)| dy \quad , \quad x \in \mathbb{R}^d \quad , \quad (2.1)$$

where the supremum is taken over all cubes with sides parallel to the coordinate axes containing  $x$ . A vector valued generalization of the classical Hardy-Littlewood maximal inequality is due to Fefferman and Stein [10].

**Theorem 2.1.** For  $1 < p < \infty$  and  $1 < \theta \leq \infty$  there exists a constant  $c > 0$ , such that

$$\left\| \left( \sum_{\ell \in I} |Mf_\ell|^\theta \right)^{1/\theta} \right\|_p \leq c \left\| \left( \sum_{\ell \in I} |f_\ell|^\theta \right)^{1/\theta} \right\|_p$$

holds for all sequences  $\{f_\ell\}_{\ell \in I}$  of locally Lebesgue-integrable functions on  $\mathbb{R}^d$ .

We require a direction-wise version of (2.1)

$$(M_i f)(x) = \sup_{s>0} \frac{1}{2s} \int_{x_i-s}^{x_i+s} |f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_d)| dt \quad , \quad x \in \mathbb{R}^d.$$

We denote the composition of these operators by  $M_e = \prod_{i \in e} M_i$ , where  $e$  is a subset of  $[d]$  and  $M_\ell M_k$  has to be interpreted as  $M_\ell \circ M_k$ . The following version of the Fefferman-Stein maximal inequality is due Stöckert [27].

**Theorem 2.2.** For  $1 < p < \infty$  and  $1 < \theta \leq \infty$  there exists a constant  $c > 0$ , such that for any  $i \in [d]$

$$\left\| \left( \sum_{\ell \in I} |M_i f_\ell|^\theta \right)^{1/\theta} \right\|_p \leq c \left\| \left( \sum_{\ell \in I} |f_\ell|^\theta \right)^{1/\theta} \right\|_p$$

holds for all sequences  $\{f_\ell\}_{\ell \in I}$  of locally Lebesgue-integrable functions on  $\mathbb{R}^d$ .

Iteration of this theorem yields a similar boundedness property for the operator  $M_{[d]}$ .

The following construction of a maximal function is due to Peetre, Fefferman and Stein. Let  $b = (b_1, \dots, b_d) > 0$ ,  $a > 0$ , and  $f \in L_1(\mathbb{R}^d)$  with  $\mathcal{F}f$  compactly supported. We define the Peetre maximal function  $P_{b,a}f$  by

$$P_{b,a}f(x) = \sup_{z \in \mathbb{R}^d} \frac{|f(x-z)|}{(1 + |b_1 z_1|)^a \cdot \dots \cdot (1 + |b_d z_d|)^a} \quad . \quad (2.2)$$

**Lemma 2.3.** Let  $\Omega \subset \mathbb{R}^d$  be a compact set. Let further  $a > 0$  and  $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ . Then there exist two constants  $c_1, c_2 > 0$  (independently of  $f$ ) such that

$$\begin{aligned} P_{(1, \dots, 1), a}(D^\alpha f)(x) &\leq c_1 P_{(1, \dots, 1), a}f(x) \\ &\leq c_2 (M_d(M_{d-1}(\dots(M_1|f|^{1/a})\dots)))^a(x) \end{aligned} \quad (2.3)$$

holds for all  $f \in L_1(\mathbb{R}^d)$  with  $\text{supp}(\mathcal{F}f) \subset \Omega$  and all  $x \in \mathbb{R}^d$ . The constants  $c_1, c_2$  depend on  $\Omega$ .



We will need the vector-valued Peetre maximal inequality which is a direct consequence of Lemma 2.3 together with Theorem 2.2.

**Theorem 2.4.** Let  $0 < p < \infty$ ,  $0 < \theta \leq \infty$  and  $a > \max\{1/p, 1/\theta\}$ . Let further  $b^\ell = (b_1^\ell, \dots, b_d^\ell) > 0$  for  $\ell \in I$  and  $\Omega = \{\Omega_\ell\}_{\ell \in I}$ , such that

$$\Omega_\ell \subset [-b_1^\ell, b_1^\ell] \times \dots \times [-b_d^\ell, b_d^\ell]$$

is compact for  $\ell \in I$ . Then there is a constant  $C > 0$  (independently of  $f$  and  $\Omega$ ) such that

$$\left\| \left( \sum_{\ell \in I} |P_{b^\ell, a} f_\ell|^\theta \right)^{1/\theta} \right\|_p \leq C \left\| \left( \sum_{\ell \in I} |f_\ell|^\theta \right)^{1/\theta} \right\|_p$$

holds for all systems  $f = \{f_\ell\}_{\ell \in I}$  with  $\text{supp}(\mathcal{F}f_\ell) \subset \Omega_\ell$ ,  $\ell \in I$ .

### 3 Function spaces of dominating mixed smoothness

In this section we introduce the function spaces under consideration, namely, the Besov and Triebel-Lizorkin spaces of dominating mixed smoothness. Note that the Sobolev spaces  $\mathbf{W}_p^s$  of mixed smoothness appear as a special case of the Triebel-Lizorkin spaces, namely  $\mathbf{W}_p^s = \mathbf{F}_{p,2}^s$  if  $1 < p < \infty$  and  $s > 0$ . There are several equivalent characterizations of these spaces, see e.g. [42]. We begin with the usual definition of the spaces using a dyadic decomposition of unity on the Fourier side, following Triebel [36]. Afterwards, we present an equivalent characterization by mixed iterated differences. To begin with, we introduce the concept of *smooth dyadic decomposition of unity*. Let  $\{\eta_j\}_{j \in \mathbb{N}_0} \subset C_0^\infty(\mathbb{R})$  such that

- (i)  $\text{supp}(\eta_0) \subset \{t: |t| \leq 2\}$ ,
- (ii)  $\text{supp}(\eta_j) \subset \{t: 2^{j-1} \leq |t| \leq 2^{j+1}\}$ ,  $j \in \mathbb{N}$ ,
- (iii) for all  $\ell \in \mathbb{N}_0$  it holds  $\sup_{t,j} 2^{j\ell} |D^\ell \eta_j(t)| \leq c_\ell < \infty$  and
- (iv)  $\sum_{j \in \mathbb{N}_0} \eta_j(t) = 1$  for all  $t \in \mathbb{R}$ .

Such a family of functions obviously exists. To approach functions spaces with dominating mixed smoothness we define the family  $\{\eta_j\}_{j \in \mathbb{N}_0^d}$  of  $d$ -variate functions as the tensor products

$$\eta_j(x) = \prod_{i=1}^d \eta_{j_i}(x_i),$$

where  $j = (j_1, \dots, j_d) \in \mathbb{N}_0^d$  and  $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ .

#### 3.1 Spaces on $\mathbb{R}^d$

Let us start with the definition of the function spaces  $\mathbf{A}_{p,\theta}^s = \mathbf{A}_{p,\theta}^s(\mathbb{R}^d)$  with  $\mathbf{A} \in \{\mathbf{B}, \mathbf{F}\}$  defined on the entire  $\mathbb{R}^d$ .

**Definition 3.1** (Besov space). Let  $0 < p, \theta \leq \infty$ ,  $s \in \mathbb{R}$ , and  $\{\eta_j\}_{j \in \mathbb{N}_0^d}$  be as above. The *Besov space of dominating mixed smoothness*  $\mathbf{B}_{p,\theta}^s = \mathbf{B}_{p,\theta}^s(\mathbb{R}^d)$  is the set of all  $f \in \mathcal{S}'(\mathbb{R}^d)$  such that

$$\|f\|_{\mathbf{B}_{p,\theta}^s} := \left( \sum_{j \in \mathbb{N}_0^d} 2^{s|j|_1 \theta} \left\| \mathcal{F}^{-1}[\eta_j \mathcal{F}f] \right\|_p^\theta \right)^{1/\theta} < \infty$$

with the usual modification for  $\theta = \infty$ .

**Definition 3.2** (Triebel-Lizorkin space). Let  $0 < p < \infty$ ,  $0 < \theta \leq \infty$ ,  $s \in \mathbb{R}$ , and  $\{\eta_j\}_{j \in \mathbb{N}_0^d}$  be as above. The *Triebel-Lizorkin space of dominating mixed smoothness*  $\mathbf{F}_{p,\theta}^s = \mathbf{F}_{p,\theta}^s(\mathbb{R}^d)$  is the set of all  $f \in \mathcal{S}'(\mathbb{R}^d)$  such that

$$\|f\|_{\mathbf{F}_{p,\theta}^s} := \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{s|j|_1 \theta} |\mathcal{F}^{-1}[\eta_j \mathcal{F}f]|^\theta \right)^{1/\theta} \right\|_p < \infty$$

with the usual modification for  $\theta = \infty$ .

**Remark 3.3.** In the special case  $\theta = 2$  and  $1 < p < \infty$  we put  $\mathbf{W}_p^s := \mathbf{F}_{p,2}^s$  which denotes the *Sobolev spaces of dominating mixed smoothness*. It is well-known (cf. [24, Chapt. 2] for  $d = 2$ ), that in case  $s \in \mathbb{N}_0$  the spaces  $\mathbf{W}_p^s$  can be equivalently normed by

$$\|f\|_{\mathbf{W}_p^s} \asymp \left( \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha|_\infty \leq s}} \|D^\alpha f\|_p^p \right)^{1/p}.$$

**Remark 3.4.** Different choices of smooth dyadic decompositions of unity  $\{\eta_j\}_{j \in \mathbb{N}_0}$  lead to equivalent (quasi-)norms.

The next lemma collects some frequently used embedding properties of the spaces.

**Lemma 3.5.** Let  $0 < p < \infty$ ,  $0 < \theta \leq \infty$ ,  $s \in \mathbb{R}$  and  $\mathbf{A} \in \{\mathbf{B}, \mathbf{F}\}$ .

(i) If  $s > \max\{1/p - 1, 0\}$  then

$$\mathbf{A}_{p,\theta}^s \hookrightarrow L_{\max\{p,1\}}(\mathbb{R}^d).$$

(ii) If  $s > 1/p$  then

$$\mathbf{A}_{p,\theta}^s \hookrightarrow C(\mathbb{R}^d).$$

*Proof.* For a proof we refer to [24, Chapt. 2]. □

Lemma 3.5(ii) gives a sufficient condition for  $\mathbf{A}_{p,\theta}^s$  to consist only of continuous functions. This is of particular importance for the analysis of cubature rules since function evaluations have to be defined reasonably.

### 3.2 Spaces on the $d$ -torus and non-periodic functions on the cube $[0, 1]^d$

Let  $\mathbb{T}^d = [0, 1]^d$  and  $f : \mathbb{T}^d \rightarrow \mathbb{C}$  be a function defined on  $\mathbb{R}^d$  and supposed to be 1-periodic in every component. We will define periodicity in a wider sense. A distribution  $f \in \mathcal{S}'(\mathbb{R})$  is a periodic distribution from  $\mathcal{S}'_\pi(\mathbb{R}^d)$  if and only if

$$f(\varphi) = f(\varphi(\cdot + k)), \quad \varphi \in \mathcal{S}(\mathbb{R}^d), \quad k \in \mathbb{Z}^d.$$

Now we define

$$\mathbf{A}_{p,\theta}^s(\mathbb{T}^d) = \{f \in \mathcal{S}'_\pi(\mathbb{R}^d) : \|f\|_{\mathbf{A}_{p,\theta}^s(\mathbb{T}^d)} < \infty\}$$

with

$$\|f\|_{\mathbf{B}_{p,\theta}^s(\mathbb{T}^d)} := \left( \sum_{j \in \mathbb{N}_0^d} 2^{s|j|_1 \theta} \|\mathcal{F}^{-1}[\eta_j \mathcal{F}f]\|_{L_p(\mathbb{T}^d)}^\theta \right)^{1/\theta}$$

and

$$\|f\|_{\mathbf{F}_{p,\theta}^s(\mathbb{T}^d)} := \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{s|j|_1 \theta} |\mathcal{F}^{-1}[\eta_j \mathcal{F}f]|^\theta \right)^{1/\theta} \right\|_{L_p(\mathbb{T}^d)},$$

respectively. Note, that we replace the  $L_p(\mathbb{R}^d)$ -norms by  $L_p(\mathbb{T}^d)$ -norms. If we denote by  $\{c_k(f)\}_{k \in \mathbb{Z}^d}$  the Fourier coefficients of  $f$ , then we could write

$$\mathcal{F}^{-1}[\eta_j \mathcal{F}f](x) = \sum_{k \in \mathbb{Z}^d} c_k(f) \eta_j(k) e^{2\pi i k x},$$

see [24, p. 150].

Next we turn to the subspace of functions which are supported in a compact set  $\Omega$ . That is, we define for  $s > \sigma_p = \max\{1/p - 1, 0\}$  the spaces

$$\mathring{\mathbf{A}}_{p,\theta}^s(\Omega) := \{f \in \mathbf{A}_{p,\theta}^s(\mathbb{R}^d) : \text{supp}(f) \subset \Omega\}. \quad (3.1)$$

Due to Lemma 3.5(i) this definition is reasonable. If  $\Omega = [0, 1]^d$  we write  $\mathring{\mathbf{A}}_{p,\theta}^s$  instead of  $\mathring{\mathbf{A}}_{p,\theta}^s([0, 1]^d)$ . The space  $\mathring{\mathbf{A}}_{p,\theta}^s$  can be interpreted as subspaces of  $\mathbf{A}_{p,\theta}^s(\mathbb{T}^d)$ . Finally, we define the spaces  $\mathbf{A}_{p,\theta}^s([0, 1]^d)$ . For large enough  $s$ , i.e.,  $s > \sigma_p$ , we define  $\mathbf{A}_{p,\theta}^s([0, 1]^d)$  as the collection of all functions  $f \in \mathbf{A}_{p,\theta}^s$  restricted to the unit cube  $[0, 1]^d$ , i.e.,

$$\mathbf{A}_{p,\theta}^s([0, 1]^d) := \{f = g|_{[0,1]^d} : g \in \mathbf{A}_{p,\theta}^s(\mathbb{R}^d)\}.$$

The (quasi-)norm on these spaces is usually given by

$$\|f\|_{\mathbf{A}_{p,\theta}^s([0,1]^d)} := \inf \{ \|g\|_{\mathbf{A}_{p,\theta}^s} : g \in \mathbf{A}_{p,\theta}^s \text{ and } g|_{[0,1]^d} = f \}.$$

### 3.3 Characterization by mixed iterated differences

In this section we will provide a characterization of the above defined spaces which is actually the classical way of defining them, see for instance Nikol'skij [20], Amanov [1], Schmeißer, Triebel [24], Temlyakov [31] and the references therein. For us it will be convenient to replace the classical moduli of continuity by so-called rectangular means of differences, see (3.3) below. This is the counterpart of the ball means of differences for isotropic spaces, see [33, Thm. 2.5.11].

Let us first recall the basic concepts. For univariate functions  $f : \mathbb{R} \rightarrow \mathbb{C}$  the  $m$ th difference operator  $\Delta_h^m$  is defined by

$$\Delta_h^m(f, t) := \sum_{j=0}^m (-1)^{m-j} \binom{m}{j} f(t + jh) \quad , \quad t \in \mathbb{R}, \quad h \in \mathbb{R}. \quad (3.2)$$

Let  $e$  be any subset of  $[d] = \{1, \dots, d\}$ . For multivariate functions  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  and  $h \in \mathbb{R}^d$  the mixed  $(m, e)$ th difference operator  $\Delta_h^{m,e}$  is defined by

$$\Delta_h^{m,e} := \prod_{i \in e} \Delta_{h_{i,i}}^m \quad \text{and} \quad \Delta_h^{m,\emptyset} = \text{Id},$$

where  $\text{Id } f = f$  and  $\Delta_{h_{i,i}}^m$  is the univariate operator applied to the  $i$ -th coordinate of  $f$  with the other variables kept fixed. Whenever we write the “product”  $\prod_{i \in e} A_i$  of operators  $A_i$  we mean the concatenation of the  $A_i$ , see also the small paragraph before Theorem 2.2 above.

Now we define the so-called rectangular means of differences by

$$\mathcal{R}_m^e(f, t, x) := \int_{[-1,1]^d} |\Delta_{(h_1 t_1, \dots, h_d t_d)}^{m,e}(f, x)| dh, \quad x \in \mathbb{R}^d, t \in (0, 1]^d. \quad (3.3)$$

For  $j \in \mathbb{N}_0^d$  we put  $e(j) = \{i : j_i \neq 0\}$  and  $e_0(j) = [d] \setminus e(j)$ . We have the following theorems.

**Theorem 3.6.** Let  $0 < p, \theta \leq \infty$  and  $s > \sigma_p$ . Let further  $m \in \mathbb{N}$  be a natural number with  $m > s$ . Then

$$\|f\|_{\mathbf{B}_{p,\theta}^s} \asymp \|f\|_{\mathbf{B}_{p,\theta}^s}^{(m)}, \quad f \in L_1(\mathbb{R}^d),$$

where

$$\|f\|_{\mathbf{B}_{p,\theta}^s}^{(m)} := \left( \sum_{j \in \mathbb{N}_0^d} 2^{s|j|_1 \theta} \|\mathcal{R}_m^{e(j)}(f, 2^{-j}, \cdot)\|_p^\theta \right)^{1/\theta}.$$

In case  $\theta = \infty$  the sum above is replaced by the supremum over  $j$ . Here  $2^{-j} := (2^{-j_1}, \dots, 2^{-j_d})$ .

**Theorem 3.7.** Let  $0 < p < \infty$ ,  $0 < \theta \leq \infty$  and  $s > \sigma_{p,\theta}$ . Let further  $m \in \mathbb{N}$  be a natural number with  $m > s$ . Then

$$\|f\|_{\mathbf{F}_{p,\theta}^s} \asymp \|f\|_{\mathbf{F}_{p,\theta}^s}^{(m)}, \quad f \in L_1(\mathbb{R}^d),$$

where

$$\|f\|_{\mathbf{F}_{p,\theta}^s}^{(m)} := \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{s|j|_1 \theta} \mathcal{R}_m^{e(j)}(f, 2^{-j}, \cdot)^\theta \right)^{1/\theta} \right\|_p.$$

Since the characterizations of the spaces  $\mathbf{B}_{p,\theta}^s$  and  $\mathbf{F}_{p,\theta}^s$  that are given in Theorems 3.6 & 3.7 are, as far as we know, not proven in the literature, we will give a proof here for the convenience of the reader.

**Proof of Theorems 3.6 & 3.7.** The statements in Theorems 3.6 & 3.7 are “discretized versions” of the characterizations given in [39]. In addition, the range of the sum  $\sum_{j \in \mathbb{N}_0^d}$  is crucial. Here, we mainly have to establish the relation  $\|f\|_{\mathbf{F}_{p,\theta}^s} \asymp \|f\|_{\mathbf{F}_{p,\theta}^s}^{(m)}$  which is done via the characterization for the spaces  $\mathbf{B}_{p,\theta}^s$  and  $\mathbf{F}_{p,\theta}^s$  via local means, see [42], [41]. Given a function  $\Psi : \mathbb{R} \rightarrow \mathbb{C}$ , we denote by  $L_\Psi \in \mathbb{N}$  the maximal number such that  $\Psi$  has vanishing moments, i.e.,

$$\int_{\mathbb{R}} t^\alpha \Psi(t) dt = 0, \quad \alpha = 0, \dots, L_\Psi.$$

Let  $\Psi_0 \in S(\mathbb{R})$  be a function satisfying the following conditions

$$\int_{\mathbb{R}} \Psi_0(t) dt \neq 0 \quad \text{and} \quad L_\Psi \geq R \quad \text{for} \quad \Psi(t) = \Psi_0(t) - \frac{1}{2} \Psi_0(t/2), \quad (3.4)$$

for some  $R \in \mathbb{N}$ . For  $k \in \mathbb{N}$  we put  $\Psi_k(t) = 2^k \Psi(2^k t)$ ,  $t \in \mathbb{R}$ . If  $j \in \mathbb{N}_0^d$  we denote

$$\Psi_j(x_1, \dots, x_d) = \Psi_{j_1}(x_1) \cdot \dots \cdot \Psi_{j_d}(x_d), \quad x \in \mathbb{R}^d.$$

We have the following proposition, see [42, Thm. 1.23] and [41].

**Proposition 3.8.** Let  $s \in \mathbb{R}$  and  $\Psi_0$  be given by (3.4) with  $R + 1 > s$ .

(i) Let  $0 < p, \theta \leq \infty$ . Then  $\mathbf{B}_{p,\theta}^s$  is the collection of all  $f \in S'(\mathbb{R}^d)$  such that

$$\|f\|_{\mathbf{B}_{p,\theta}^s} = \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \|\Psi_j * f\|_p^\theta \right)^{1/\theta} < \infty.$$

(ii) Let  $0 < p < \infty$  and  $0 < \theta \leq \infty$ . Then  $\mathbf{F}_{p,\theta}^s$  is the collection of all  $f \in S'(\mathbb{R}^d)$  such that

$$\|f\|_{\mathbf{F}_{p,\theta}^s} = \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} |\Psi_j * f|^\theta \right)^{1/\theta} \right\|_p < \infty.$$

Now we are in a position to prove Theorem 3.7. The proof of Theorem 3.6 is similar but technically less involved.

*Proof of Theorem 3.7. Step 1.* We prove the inequality

$$\|f\|_{\mathbf{F}_{p,\theta}^s}^{(m)} \lesssim \|f\|_{\mathbf{F}_{p,\theta}^s}. \quad (3.5)$$

This inequality is a consequence of [39, Thm. 3.4.1]. Indeed, from the equivalent norm

$$\|f\|_{\mathbf{F}_{p,\theta}^s} \asymp \|f\|_p + \sum_{e \subset [d], e \neq \emptyset} \left\| \left( \int_{(0,\infty)^{|e|}} \left( \prod_{i \in e} t_i^{-s\theta} \right) \mathcal{R}_m^e(f, t, \cdot)^\theta \prod_{i \in e} \frac{dt_i}{t_i} \right)^{1/\theta} \right\|_p,$$

see [39, Thm. 3.4.1], we have

$$\|f\|_{\mathbf{F}_{p,\theta}^s} \gtrsim \|f\|_p + \sum_{e \subset [d], e \neq \emptyset} \left\| \left( \int_{(0,1)^{|e|}} \left( \prod_{i \in e} t_i^{-s\theta} \right) \mathcal{R}_m^e(f, t, \cdot)^\theta \prod_{i \in e} \frac{dt_i}{t_i} \right)^{1/\theta} \right\|_p.$$

Discretizing the right-hand side of the above inequality we obtain

$$\begin{aligned} & \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \mathcal{R}_m^{e(j)}(f, 2^{-j}, \cdot)^\theta \right)^{1/\theta} \right\|_p \\ & \lesssim \|f\|_p + \sum_{e \subset [d], e \neq \emptyset} \left\| \left( \int_{(0,1)^{|e|}} \left( \prod_{i \in e} t_i^{-s\theta} \right) \mathcal{R}_m^e(f, t, \cdot)^\theta \prod_{i \in e} \frac{dt_i}{t_i} \right)^{1/\theta} \right\|_p, \end{aligned}$$

which implies (3.5).

*Step 2.* We prove the reverse inequality in (3.5). This time we rely on the intrinsic characterization of  $\mathbf{F}_{p,\theta}^s$  via local means, see Proposition 3.8.

*Substep 2.1.* Preparation. According to [34, 3.3.1, 3.3.2], for  $m \in \mathbb{N}$  there exist function  $\psi_m(t)$  with

$$\text{supp}(\psi) \subset \left( -\frac{1}{m}, \frac{1}{m} \right) \quad \text{and} \quad \int_{\mathbb{R}} \psi_m(t) dt = 1$$

such that

$$\Psi_0(t) = \frac{(-1)^{m+1}}{m!} \sum_{u=1}^m \sum_{v=1}^m (-1)^{u+v} \binom{m}{u} \binom{m}{v} v^m (uv)^{-1} \psi\left(\frac{t}{uv}\right), \quad t \in \mathbb{R},$$

satisfies (3.4) with  $R = m$ . Putting

$$\eta(t) = \frac{(-1)^{m+1}}{m!} \sum_{v=1}^m (-1)^{m-v} \binom{m}{v} v^{m-v-1} \psi\left(\frac{t}{v}\right), \quad t \in \mathbb{R},$$

then we have  $\text{supp}(\eta) \in (-1, 1)$  and moreover

$$\Psi_0(t) = \sum_{u=1}^m (-1)^{m-u} \binom{m}{u} u^{-1} \eta\left(\frac{t}{u}\right).$$

A simple computation gives for a univariate function  $g$

$$\begin{aligned} \Psi_0 * g(t) &= \int_{\mathbb{R}} \Psi_0(h) g(t-h) dh \\ &= \int_{\mathbb{R}} \eta(-h) \sum_{u=1}^m (-1)^{m-u} \binom{m}{u} g(t+uh) dh \\ &= \int_{-1}^1 \eta(-h) [\Delta_h^m g(t) - (-1)^m g(t)] dh \end{aligned} \quad (3.6)$$

and

$$\begin{aligned} \Psi_k * g(t) &= \int_{\mathbb{R}} \Psi(h) g(t-2^{-k}h) dh \\ &= \int_{\mathbb{R}} \left[ \eta(-h) - \frac{1}{2} \eta\left(-\frac{h}{2}\right) \right] \Delta_{2^{-k}h}^m g(t) dh \\ &= \int_{-1}^1 \eta(-h) [\Delta_{2^{-k}h}^m - \Delta_{2^{-k+1}h}^m] g(t) dh. \end{aligned} \quad (3.7)$$

*Substep 2.2.* We define the function  $\eta(\cdot)$  on  $\mathbb{R}^d$  as  $d$ -fold tensor product, i.e.,

$$\eta(x) = \eta(x_1) \cdot \dots \cdot \eta(x_d), \quad x = (x_1, \dots, x_d) \in \mathbb{R}^d.$$

From (3.6) and (3.7) we have for  $j \in \mathbb{N}_0^d$

$$\Psi_j * f(x) = \int_{[-1,1]^d} \eta(-h) \left[ \prod_{i \in e(j)} (\Delta_{2^{-j_i}h_{i,i}}^m - \Delta_{2^{-j_i+1}h_{i,i}}^m) \prod_{i \in e_0(j)} (\Delta_{h_{i,i}}^m - (-1)^m \text{Id}_i) \right] f(x) dh.$$

Then we obtain

$$|\Psi_j * f(x)| \leq \sum_{u \in \{0,1\}^d} \mathcal{R}_m^{e(j+u)}(f, 2^{-j+u}, x)$$

which leads to

$$\begin{aligned} \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{s|j|_1 \theta} |\Psi_j * f|^\theta \right)^{1/\theta} \right\|_p &\lesssim \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{s|j|_1 \theta} \sum_{u \in \{0,1\}^d} \mathcal{R}_m^{e(j+u)}(f, 2^{-j+u}, \cdot)^\theta \right)^{1/\theta} \right\|_p \\ &\lesssim \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{s|j|_1 \theta} \mathcal{R}_m^{e(j)}(f, 2^{-j}, \cdot)^\theta \right)^{1/\theta} \right\|_p. \end{aligned}$$

In view of Proposition 3.8 we finish the proof. □

## 4 Change of variables

Let  $\varphi \in C_0^k(\mathbb{R})$  such that  $\text{supp}(\varphi) \subset [0, 1]$ ,  $\int_0^1 \varphi(t) dt = 1$ ,  $\varphi(t) > 0$  on  $(0, 1)$  and  $\varphi^{(k)}$  has only finitely many zeros in  $[0, 1]$ . Let further

$$\psi(t) := \int_{-\infty}^t \varphi(\xi) d\xi, \quad t \in \mathbb{R},$$

then  $\psi'(t) = \varphi(t)$ ,  $t \in \mathbb{R}$ . We study the boundedness of the “change of variable” operator

$$T_d^\psi: f(x) \mapsto \left( \prod_{i=1}^d \psi'(x_i) \right) \cdot f(\psi(x_1), \dots, \psi(x_d)), \quad f \in L_1(\mathbb{R}^d), \quad x \in \mathbb{R}^d,$$

in the situations

$$T_d^\psi: \mathbf{B}_{p,\theta}^s(\mathbb{R}^d) \rightarrow \mathbf{B}_{p,\theta}^s(\mathbb{R}^d) \quad \text{and} \quad T_d^\psi: \mathbf{F}_{p,\theta}^s(\mathbb{R}^d) \rightarrow \mathbf{F}_{p,\theta}^s(\mathbb{R}^d).$$

The following theorems represent the main results of the paper.

**Theorem 4.1.** Let  $1 \leq p \leq \infty$ ,  $0 < \theta \leq \infty$ ,  $s > 0$  and  $\varphi \in C_0^k(\mathbb{R})$  as above with  $k > [s] + 2$  if  $p > 1$  and  $k > [s] + 3$  if  $p = 1$ . Then

$$\|T_d^\psi f\|_{\mathbf{B}_{p,\theta}^s} \lesssim_\psi \|f\|_{\mathbf{B}_{p,\theta}^s}, \quad f \in \mathbf{B}_{p,\theta}^s(\mathbb{R}^d).$$

**Theorem 4.2.** Let  $1 < p < \infty$ ,  $1 < \theta \leq \infty$ ,  $s > 0$  and  $\varphi \in C_0^k(\mathbb{R})$  as above with  $k > [s] + 2$ . Then

$$\|T_d^\psi f\|_{\mathbf{F}_{p,\theta}^s} \lesssim_\psi \|f\|_{\mathbf{F}_{p,\theta}^s}, \quad f \in \mathbf{F}_{p,\theta}^s(\mathbb{R}^d).$$

**Remark 4.3.** (i) Observe that the smoothness  $k$  of kernel  $\varphi$  in the  $\mathbf{F}$ -case does not have to grow to infinity when  $p$  tends to 1. This result has to be compared with the mentioned result of Templyakov, see (1.8) and [32, page 237].

(ii) With the technique used in this paper we were not able to give a result in the case  $p < 1$  for both,  $\mathbf{B}$  and  $\mathbf{F}$ -spaces, not even in the univariate situation, see Lemma 4.4 below. Note, that also the case  $p = 1$  is open in the  $\mathbf{F}$ -case. However, we strongly conjecture that we can prove boundedness of the change of variable mapping if  $1/2 < p \leq 1$  in both cases.

For the convenience of the reader we will first comment on a corresponding univariate result.

**Lemma 4.4.** Let  $1 < p < \infty$ ,  $1 < \theta \leq \infty$ ,  $s > 0$  and  $\varphi \in C_0^k(\mathbb{R})$  as above with  $k > [s] + 2$ . Then

$$\|T_1^\psi f\|_{F_{p,\theta}^s} \lesssim_\psi \|f\|_{F_{p,\theta}^s}, \quad f \in F_{p,\theta}^s(\mathbb{R}).$$

Before we come to the proof of the univariate boundedness results we need some preparation. It will turn out that difference characterizations of the function spaces under consideration is a suitable approach to analyze the respective operators.

## 4.1 Proof of Lemma 4.4

Let us first prove a technical lemma on the boundedness of quotients of derivatives of  $\varphi$ . Such terms naturally appear if one bounds the  $L_p$ -norm of  $T_1^\psi f$ . For the particular choice  $\varphi = \psi'_k$ , with  $\psi_k$  from (1.6), the lemma below was proved in [31, p. 238].

**Lemma 4.5.** Let  $1 < p \leq \infty$  and  $m \in \mathbb{N}_0$ . Further let  $\varphi \in C_0^k(\mathbb{R})$  with  $\text{supp}(\varphi) \subset [0, 1]$ ,  $\varphi > 0$  on  $(0, 1)$  and  $k > \frac{mp}{p-1} + 1$ . Let us further assume that the  $k$ th derivative  $\varphi^{(k)}$  has only finitely many zeros in  $[0, 1]$ . Then we have

$$\frac{|\varphi^{(n)}(t)|}{|\varphi(t)|^{1/p}} \in L_\infty([0, 1])$$

for all  $0 \leq n \leq m$ .

*Proof.* It is enough to prove the result for  $n = m$ . If  $p = \infty$  the result is obvious. Hence, assume  $1 < p < \infty$ . We put  $\ell = k - 1$ . Using Taylor's theorem and the fact that  $\varphi^{(i)}(0) = 0$  for all  $i = 0, \dots, k$ , we obtain

$$\varphi(t) = \frac{1}{\ell!} \int_0^t \varphi^{(\ell+1)}(\xi) (t - \xi)^\ell d\xi$$

as well as

$$\varphi^{(m)}(t) = \frac{1}{(\ell - m)!} \int_0^t \varphi^{(\ell+1)}(\xi) (t - \xi)^{\ell-m} d\xi$$

for all  $t \in [0, 1]$ . Since  $\varphi^{(\ell+1)}$  has only finitely many zeros in  $[0, 1]$  there exists an  $\varepsilon > 0$  such that  $\varphi^{(\ell+1)}(t) > 0$  for  $t \in (0, \varepsilon)$ . This shows with  $p' = p/(p-1)$  that

$$\begin{aligned} \frac{|\varphi^{(m)}(t)|}{|\varphi(t)|^{1/p}} &\leq \frac{\ell!}{(\ell - m)!} \frac{t^{1/p'} \left( \int_0^t (\varphi^{(\ell+1)}(\xi))^p (t - \xi)^{p(\ell-m)} d\xi \right)^{1/p}}{\left( \int_0^t \varphi^{(\ell+1)}(\xi) (t - \xi)^\ell d\xi \right)^{1/p}} \\ &\lesssim_k \sup_{\xi \in (0, \varepsilon)} \left( |\varphi^{(\ell+1)}(\xi)|^{p-1} |t - \xi|^{p(\ell-m)-\ell} \right)^{1/p} \frac{\left( \int_0^t \varphi^{(\ell+1)}(\xi) (t - \xi)^\ell d\xi \right)^{1/p}}{\left( \int_0^t \varphi^{(\ell+1)}(\xi) (t - \xi)^\ell d\xi \right)^{1/p}} \\ &\lesssim_k 1 \end{aligned}$$

for  $t \in (0, \varepsilon)$ . In the last inequality we used  $p > 1$  and  $\ell > \frac{pm}{p-1}$ . The same arguments work also for  $(1 - \varepsilon_0, 1)$  with some  $\varepsilon_0 > 0$ . The quotient is uniformly bounded in  $[\varepsilon, 1 - \varepsilon_0]$  since  $\varphi(t) \geq c > 0$  for  $t \in [\varepsilon, 1 - \varepsilon_0]$ . □

It is easily seen that Lemma 4.5 is not true for  $p = 1$ . We immediately obtain the following corollary. Note the relaxed conditions on  $p$  and  $k$ .

**Corollary 4.6.** Let  $1 \leq p \leq \infty$  and  $m \in \mathbb{N}_0$ . Further let  $\varphi \in C_0^k(\mathbb{R})$  with  $\text{supp}(\varphi) \subset [0, 1]$ ,  $\varphi > 0$  on  $(0, 1)$  and  $k > m + 1$ . Let us further assume that the  $k$ th derivative  $\varphi^{(k)}$  has only finitely many zeros in  $[0, 1]$ . Then we have

$$\frac{|\varphi^{(r)}(t)\varphi^{(\alpha)}(t)|}{|\varphi(t)|^{1/p}} \in L_\infty([0, 1])$$

for all  $0 \leq r, \alpha \leq m$  with  $r + \alpha \leq m$ .



*Proof.* Obviously, it is enough to prove the statement for  $p = 1$ , since

$$\frac{|\varphi^{(r)}(t)\varphi^{(\alpha)}(t)|}{|\varphi(t)|^{1/p}} \lesssim \frac{|\varphi^{(r)}(t)\varphi^{(\alpha)}(t)|}{|\varphi(t)|}$$

due to the boundedness of  $\varphi$ . We write

$$\frac{|\varphi^{(r)}(t)\varphi^{(\alpha)}(t)|}{|\varphi(t)|} = \frac{|\varphi^{(r)}(t)|}{|\varphi(t)|^{1/p_1}} \cdot \frac{|\varphi^{(\alpha)}(t)|}{|\varphi(t)|^{1/p_2}}$$

for some  $p_1, p_2$  with  $1/p_1 + 1/p_2 = 1$ . Using Lemma 4.5 we know that we need  $k > \frac{rp_1}{p_1-1} + 1$  for the boundedness of the first factor and  $k > \frac{\alpha p_2}{p_2-1} + 1$  for the boundedness of the second.

First note that the statement of the corollary is trivial for  $r = \alpha = 0$ . If only one of them is zero, say  $r = 0$ , then we set  $p_1 = 1$  and  $p_2 = \infty$ ; again the statement is obvious. Hence, we assume now that  $r \neq 0$  and  $\alpha \neq 0$  and we choose  $p_1 = (r + \alpha)/\alpha$  and  $p_2 = (r + \alpha)/r$ .

This leads to the restriction  $k > r + \alpha + 1$  in both cases. Since we want a result for all  $r + \alpha \leq m$  we impose the restriction  $k > m + 1$ . This finishes the proof.  $\square$

Now we are in position to prove Lemma 4.4.

**Proof of Lemma 4.4.** *Step 1.* Let  $\eta_j$  denote a smooth dyadic compactly supported decomposition of unity (for instance the one given at the beginning of Subsection 3.1). In particular,  $1 = \sum_{j \in \mathbb{N}_0} \eta_j(t)$  for all  $t \in \mathbb{R}$  and  $\eta_j(t) = 0$  if  $|t| > 2^{j+1}$ . We decompose  $f$  as follows

$$f(t) = \sum_{j \in \mathbb{Z}} f_j(t) := \sum_{j \in \mathbb{Z}} \mathcal{F}^{-1}[\eta_j \mathcal{F}f](t), \quad (4.1)$$

where we put  $f_j \equiv 0$  in case  $j < 0$ . Note, that with this definition, we can clearly write also  $f = \sum_{\ell \in \mathbb{Z}} f_{j+\ell}$  for every fixed  $j \in \mathbb{N}_0$ . Using the characterization by rectangular means of differences of  $F_{p,\theta}^s$ , see Theorem 3.7, and the decomposition (4.1) we have

$$\begin{aligned} \|T_1^\psi f\|_{F_{p,\theta}^s}^{(m)} &\leq \sum_{\ell \in \mathbb{Z}} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(T_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{1/\theta} \right\|_p \\ &= \sum_{\ell < 0} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(T_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{1/\theta} \right\|_p \\ &\quad + \sum_{\ell \geq 0} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(T_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{1/\theta} \right\|_p. \end{aligned} \quad (4.2)$$

Here we choose  $m = \lfloor s \rfloor + 1$ .

*Step 2.* First we recall the convolution representation of the  $m$ th order difference. Let  $M_m(\cdot)$  is univariate B-spline of degree  $m$  which has knots at the points  $\{0, 1, \dots, m\}$ , i.e.,

$$M_m = \chi_{[0,1]} * \dots * \chi_{[0,1]}, \quad m \text{ times},$$

where  $\chi_{[0,1]}$  denotes the characteristic function of  $[0, 1]$ . Note, that  $M_m(\cdot)$  is bounded and  $\text{supp}(M_m) \subset [0, m]$ . In case  $h > 0$  we have

$$\Delta_h^m(T_1^\psi f_{j+\ell}, t) = h^{m-1} \int_0^{hm} (T_1^\psi f_{j+\ell})^{(m)}(t + \xi) M_m(h^{-1}\xi) d\xi,$$

see [5, page 45]. If  $h < 0$  we use B-spline with knots  $\{-m, \dots, 0\}$  and get the similar formula. Consequently, we obtain for  $|h| < 2^{-j}$

$$|\Delta_h^m(T_1^\psi f_{j+\ell}, t)| \lesssim |h|^m M[(T_1^\psi f_{j+\ell})^{(m)}](t) \leq 2^{-jm} M[(T_1^\psi f_{j+\ell})^{(m)}](t), \quad (4.3)$$

since  $M_m(h^{-1} \cdot)$  is bounded. This leads to

$$\begin{aligned} \sum_{\ell < 0} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(T_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{1/\theta} \right\|_p \\ \lesssim \sum_{\ell < 0} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{(s-m)j\theta} (M[(T_1^\psi f_{j+\ell})^{(m)}](\cdot))^\theta \right)^{1/\theta} \right\|_p \\ \lesssim \sum_{\ell < 0} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{(s-m)j\theta} |(T_1^\psi f_{j+\ell})^{(m)}(\cdot)|^\theta \right)^{1/\theta} \right\|_p. \end{aligned}$$

Where in the second inequality we have applied Theorem 2.1. Now we continue with

$$(T_1^\psi f_{j+\ell})^{(m)}(t) = [\varphi f_{j+\ell}(\psi)]^{(m)}(t) = \sum_{\alpha=0}^m C_\alpha \varphi^{(\alpha)}(t) [f_{j+\ell}(\psi)]^{(m-\alpha)}(t),$$

for some  $C_\alpha$ . We distinguish two cases. First let us deal with  $\alpha = m$ . We choose the number  $a$  such that  $\max\{1/p, 1/\theta\} < a \leq m$ . This is possible since  $\max\{1/p, 1/\theta\} < 1$  and  $m \geq 1$ . Then, if  $t \in [0, 1]$  we have

$$\begin{aligned} |\varphi^{(m)}(t) f_{j+\ell}(\psi(t))| &\lesssim \sup_{\xi \in [0,1]} |f_{j+\ell}(\xi)| = \sup_{\xi \in [0,1]} \frac{|f_{j+\ell}(\xi)|}{(1 + 2^{j+\ell}|t - \xi|)^a} (1 + 2^{j+\ell}|t - \xi|)^a \\ &\lesssim 2^{(\ell+j)a} \sup_{\xi \in [0,1]} \frac{|f_{j+\ell}(\xi)|}{(1 + 2^{j+\ell}|t - \xi|)^a} \\ &\lesssim 2^{(\ell+j)m} \sup_{\xi \in \mathbb{R}} \frac{|f_{j+\ell}(\xi)|}{(1 + 2^{j+\ell}|t - \xi|)^a} \\ &= 2^{(\ell+j)m} P_{j+\ell,a}(f_{j+\ell}, t). \end{aligned} \quad (4.4)$$

If  $t \notin [0, 1]$  the inequality (4.4) is obviously satisfied since  $\text{supp}(\varphi) \subset [0, 1]$ . We consider the summand

$$\begin{aligned} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{(s-m)j\theta} |\varphi^{(m)}(\cdot) f_{j+\ell}(\psi(\cdot))|^\theta \right)^{1/\theta} \right\|_p &\lesssim 2^{(m-s)\ell} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{(\ell+j)s\theta} P_{j+\ell,a}(f_{j+\ell}, \cdot)^\theta \right)^{1/\theta} \right\|_p \\ &\lesssim 2^{(m-s)\ell} \|f\|_{F_{p,\theta}^s}. \end{aligned} \quad (4.5)$$

In the last step we used Theorem 2.4. In case  $\alpha < m$  we have

$$\varphi^{(\alpha)}(t) [f_{j+\ell}(\psi)]^{(m-\alpha)}(t) = \sum_{1 \leq \gamma \leq m-\alpha} f_{j+\ell}^{(\gamma)}(\xi)|_{\xi=\psi(t)} \cdot \varphi^{(\alpha)}(t) \cdot Q_\gamma(t),$$

where  $Q_\gamma(t)$  is the sum of the forms  $(\varphi^{(\beta_1)}(t))^{n_1} (\varphi^{(\beta_2)}(t))^{n_2}$  with some  $\beta_i, n_i \in \mathbb{N}_0$ ,  $i = 1, 2$ , satisfied  $\beta_1 n_1 + \beta_2 n_2 = \gamma$ . Note that the highest order derivative of  $\varphi(t)$  in  $Q_\gamma(t)$  is  $m - \alpha$ . Since  $\varphi \in C_0^k(\mathbb{R})$ ,  $k > [s] + 2 = m + 1$ , Corollary 4.6 yields that

$$\frac{|\varphi^{(r)}(t) \varphi^{(\alpha)}(t)|}{|\varphi(t)|^{1/p}} < C, \quad (4.6)$$

for all  $r, \alpha$  with  $r + \alpha \leq m$ , which leads to  $|\varphi^{(\alpha)}(t)Q_\gamma(t)| \lesssim |\varphi(t)|^{1/p}$ . Consequently,

$$|\varphi^{(\alpha)}(t)[f_{j+\ell}(\psi)]^{(m-\alpha)}(t)| \lesssim \sum_{1 \leq \gamma \leq m-\alpha} |f_{j+\ell}^{(\gamma)}(\xi)|_{\xi=\psi(t)} \cdot |\varphi(t)|^{1/p}.$$

We consider the term

$$\begin{aligned} F_{\alpha,\ell} &:= \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{(s-m)j\theta} |\varphi^{(\alpha)}[f_{j+\ell}(\psi)]^{(m-\alpha)}|^\theta \right)^{1/\theta} \right\|_p \\ &\lesssim \sum_{1 \leq \gamma \leq m-\alpha} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{(s-m)j\theta} |f_{j+\ell}^{(\gamma)}(\xi)|_{\xi=\psi(\cdot)} \cdot \varphi^{1/p} |^\theta \right)^{1/\theta} \right\|_p \\ &\lesssim \sum_{1 \leq \gamma \leq m-\alpha} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{(s-m)j\theta} |f_{j+\ell}^{(\gamma)}|^\theta \right)^{1/\theta} \right\|_p. \end{aligned} \quad (4.7)$$

By Lemma 2.3 and a homogeneity argument we have for any  $a > 0$

$$|f_{j+\ell}^{(\gamma)}(t)| \leq \sup_{\xi \in \mathbb{R}} \frac{|f_{j+\ell}^{(\gamma)}(\xi)|}{(1 + 2^{j+\ell}|t - \xi|)^a} \lesssim 2^{(j+\ell)m} \sup_{\xi \in \mathbb{R}} \frac{|f_{j+\ell}(\xi)|}{(1 + 2^{j+\ell}|t - \xi|)^a} \quad (4.8)$$

if  $1 \leq \gamma \leq m$ . Putting this into (4.7) with  $a > \max\{1/p, 1/\theta\}$  and using Theorem 2.4 afterwards yields

$$\begin{aligned} F_{\alpha,\ell} &\lesssim \sum_{1 \leq \gamma \leq m-\alpha} 2^{\ell(m-s)} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{(j+\ell)s\theta} |P_{j+\ell,a}(f_{j+\ell}, \cdot)|^\theta \right)^{1/\theta} \right\|_p \\ &\lesssim 2^{\ell(m-s)} \|f\|_{F_{p,\theta}^s}. \end{aligned} \quad (4.9)$$

From (4.5), (4.7), and (4.9) we obtain

$$\begin{aligned} \sum_{\ell < 0} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(T_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{1/\theta} \right\|_p &\lesssim \sum_{\ell < 0} 2^{\ell(m-s)} \|f\|_{F_{p,\theta}^s} \\ &\lesssim \|f\|_{F_{p,\theta}^s}. \end{aligned} \quad (4.10)$$

*Step 3.* It remains to deal with  $\sum_{\ell \geq 0}$ . We expand the difference and estimate

$$\begin{aligned} \sum_{\ell \geq 0} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(T_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{1/\theta} \right\|_p \\ \lesssim \sum_{\ell \geq 0} \sum_{i=0}^m \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |T_1^\psi f_{j+\ell}(\cdot + ih)| dh \right]^\theta \right)^{1/\theta} \right\|_p. \end{aligned} \quad (4.11)$$

We consider the term

$$\begin{aligned} A_{i,\ell} &= \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |(T_1^\psi f_{j+\ell})(\cdot + ih)| dh \right]^\theta \right)^{1/\theta} \right\|_p \\ &\lesssim \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} (M[T_1^\psi f_{j+\ell}](\cdot))^\theta \right)^{1/\theta} \right\|_p. \end{aligned} \quad (4.12)$$

Here again we use the Hardy-Littlewood maximal function. From Theorem 2.1 again we get

$$\begin{aligned}
A_{i,\ell} &\lesssim \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} |(T_1^\psi f_{j+\ell})|^\theta \right)^{1/\theta} \right\|_p \\
&\lesssim 2^{-\ell s} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{(j+\ell)s\theta} |\varphi(\cdot) f_{j+\ell}(\psi(\cdot))|^\theta \right)^{1/\theta} \right\|_p \\
&\lesssim 2^{-\ell s} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{(j+\ell)s\theta} |f_{j+\ell}|^\theta \right)^{1/\theta} \right\|_p \\
&\lesssim 2^{-\ell s} \|f\|_{F_{p,\theta}^s}. \tag{4.13}
\end{aligned}$$

We performed a change of variable in the second step and used  $p > 1$  and  $\varphi \in C_0^k(\mathbb{R})$ . With (4.11), (4.12), and (4.13) we obtain

$$\begin{aligned}
\sum_{\ell \geq 0} \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(T_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{1/\theta} \right\|_p \\
\lesssim \sum_{i=1}^m \sum_{\ell \geq 0} A_{i,\ell} \lesssim \sum_{\ell \geq 0} 2^{-\ell s} \|f\|_{F_{p,\theta}^s} \lesssim \|f\|_{F_{p,\theta}^s}, \tag{4.14}
\end{aligned}$$

where we used  $s > 0$  in the last inequality. Altogether we obtain from (4.10) and (4.14)

$$\left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(T_1^\psi f, \cdot)| dh \right]^\theta \right)^{1/\theta} \right\|_p \lesssim \|f\|_{F_{p,\theta}^s}.$$

In view of (4.2) we conclude that

$$\|T_1^\psi f\|_{F_{p,\theta}^s}^{(m)} \lesssim \|f\|_{F_{p,\theta}^s},$$

which finishes the proof. □

## 4.2 Proof of Theorems 4.1 and 4.2

**Proof of Theorem 4.2.** *Step 1.* We aim at adapting the proof of Lemma 4.4 to the multivariate mixed situation by applying the arguments direction-wise. The proper tool will be the characterization by rectangular means of differences, see Theorem 3.7. Let  $\{\eta_j\}_{j \in \mathbb{N}_0^d}$  be a tensorized compactly supported smooth dyadic decomposition of unity and

$$f_j := \mathcal{F}^{-1}[\eta_j \mathcal{F} f] \quad , \quad j \in \mathbb{Z}^d, \tag{4.15}$$

where we put  $f_j \equiv 0$  if  $j \notin \mathbb{N}_0^d$ . Analogously, to (4.2) we estimate the quantity

$$\begin{aligned}
&\left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \mathcal{R}_m^{e(j)}(T_d^\psi f, 2^{-j}, \cdot)^\theta \right)^{1/\theta} \right\|_p \\
&\leq \sum_{\ell \in \mathbb{Z}^d} \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \mathcal{R}_m^{e(j)}(T_d^\psi f_{j+\ell}, 2^{-j}, \cdot)^\theta \right)^{1/\theta} \right\|_p, \tag{4.16}
\end{aligned}$$

where  $2^{-j} = (2^{-j_1}, \dots, 2^{-j_d})$ . Again we choose  $m = \lfloor s \rfloor + 1$ . Let us introduce some further notation first. If  $e \subset [d]$  and  $j \in \mathbb{N}_0^d$  we denote

$$\Gamma_j^e = \bigtimes_{i \in e} [-2^{-j_i}, 2^{-j_i}]. \quad (4.17)$$

For each  $\ell \in \mathbb{Z}^d$  we denote

$$e_1(j, \ell) = e(j) \cap \{i : \ell_i \geq 0\} \quad \text{and} \quad e_2(j, \ell) = e(j) \cap \{i : \ell_i < 0\}. \quad (4.18)$$

For simplicity we put

$$F_\ell := \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \mathcal{R}_m^{e(j)}(T_d^\psi f_{j+\ell}, 2^{-j}, \cdot)^\theta \right)^{1/\theta} \right\|_p.$$

*Step 2.* Estimation of  $F_\ell$ . We have

$$\begin{aligned} \Delta_h^{m, e(j)}(T_d^\psi f_{j+\ell}, x) &= \Delta_h^{m, e_1(j, \ell)} \circ \Delta_h^{m, e_2(j, \ell)}(T_d^\psi f_{j+\ell}, x) \\ &= \sum_k \Delta_h^{m, e_2(j, \ell)}(T_d^\psi f_{j+\ell}, x + (k \cdot h)) \end{aligned} \quad (4.19)$$

where the sum on the right-hand side is taken over all  $k \in \mathbb{N}_0^d$  such that  $0 \leq k_i \leq m$  if  $i \in e_1(j, \ell)$  and  $k_i = 0$  otherwise. In addition, we put  $(k \cdot h) = (k_1 h_1, \dots, k_d h_d)$ . Using the convolution representation of the  $m$ th order difference in  $e_2(j, \ell)$  we have the inequality

$$\begin{aligned} &|\Delta_h^{m, e_2(j, \ell)}(T_d^\psi f_{j+\ell}, x + (k \cdot h))| \\ &\lesssim \left( \prod_{i \in e_2(j, \ell)} 2^{-j_i m} \right) M_{e_2(j, \ell)}[D^{(m_{e_2(j, \ell)})}(T_d^\psi f_{j+\ell})](x + (k \cdot h)), \end{aligned}$$

see (4.3). Here  $m_{e_2(j, \ell)} = (m_1, \dots, m_d)$  where  $m_i = m$  if  $i \in e_2(j, \ell)$  otherwise  $m_i = 0$ . This leads to

$$\begin{aligned} &\mathcal{R}_m^{e(j)}(T_d^\psi f_{j+\ell}, 2^{-j}, x) \\ &\lesssim \sum_k \int_{\Gamma_j^{e(j)}} |\Delta_h^{m, e_2(j, \ell)}(T_d^\psi f_{j+\ell}, x + (k \cdot h))| \prod_{i \in e(j)} 2^{j_i} dh_i \\ &\lesssim \sum_k \int_{\Gamma_j^{e(j)}} \left( \prod_{i \in e_2(j, \ell)} 2^{-j_i m} \right) M_{e_2(j, \ell)}[D^{(m_{e_2(j, \ell)})}(T_d^\psi f_{j+\ell})](x + (k \cdot h)) \prod_{i \in e(j)} 2^{j_i} dh_i. \end{aligned}$$

Now we estimate above by the Hardy-Littlewood maximal function with the components in  $e_1(j, \ell)$  and obtain

$$\begin{aligned} \mathcal{R}_m^{e(j)}(T_d^\psi f_{j+\ell}, 2^{-j}, x) &\lesssim \left( \prod_{i \in e_2(j, \ell)} 2^{-j_i m} \right) M_{e_1(j, \ell)}[M_{e_2(j, \ell)}[D^{(m_{e_2(j, \ell)})}(T_d^\psi f_{j+\ell})]](x) \\ &\lesssim \left( \prod_{i \in e_2(j, \ell)} 2^{-j_i m} \right) M_{[d]}[D^{(m_{e_2(j, \ell)})}(T_d^\psi f_{j+\ell})](x). \end{aligned}$$

Plugging this into  $F_\ell$ , Theorem 2.1 yields

$$F_\ell \lesssim \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \left( \prod_{i \in e_2(j, \ell)} 2^{-j_i m} \right) |D^{(m_{e_2(j, \ell)})}(T_d^\psi f_{j+\ell})(\cdot)|^\theta \right)^{1/\theta} \right\|_p.$$

Using Leibniz's formula we estimate

$$\begin{aligned}
|D^{(m_{e_2(j,\ell)})}(T_d^\psi f_{j+\ell})(x)| &= \left| D^{(m_{e_2(j,\ell)})} \left( \left[ \bigotimes_{i=1}^d \varphi \right] f_{j+\ell}(\psi, \dots, \psi) \right)(x) \right| \\
&\lesssim \sum_{\alpha} \left| D^{(\alpha)} \left[ \bigotimes_{i=1}^d \varphi \right](x) D^{(\beta)} [f_{j+\ell}(\psi, \dots, \psi)](x) \right|.
\end{aligned} \tag{4.20}$$

Here the sum on the right-hand side is taken over all  $\alpha \in \mathbb{N}_0^d$  with  $0 \leq \alpha_i \leq m$  if  $i \in e_2(j, \ell)$ , otherwise  $\alpha_i = 0$  and  $\beta = m_{e_2(j,\ell)} - \alpha$ . We continue estimation by dividing  $e_2(j, \ell) = e_2^1(j, \ell) \cup e_2^2(j, \ell)$  where  $\alpha_i < m$  if  $i \in e_2^1(j, \ell)$  and  $\alpha_i = m$  if  $i \in e_2^2(j, \ell)$ . We have

$$\begin{aligned}
&D^{(\alpha)} \left[ \bigotimes_{i=1}^d \varphi \right](x) D^{(\beta)} [f_{j+\ell}(\psi, \dots, \psi)](x) \\
&= \left( \prod_{i \in e_0(j) \cup e_1(j,\ell)} \varphi(x_i) \right) \left( \prod_{i \in e_2^2(j,\ell)} \varphi^{(m)}(x_i) \right) \left( \prod_{i \in e_2^1(j,\ell)} \varphi^{(\alpha_i)}(x_i) \right) D^{(\beta)} [f_{j+\ell}(\psi, \dots, \psi)](x).
\end{aligned}$$

Now the condition  $\varphi \in C_0^k(\mathbb{R})$ ,  $k > m + 1$ , implies that

$$\begin{aligned}
&\left| \left( \prod_{i \in e_2^1(j,\ell)} \varphi^{(\alpha_i)}(x_i) \right) D^{(\beta)} [f_{j+\ell}(\psi, \dots, \psi)](x) \right| \\
&\lesssim \left( \prod_{i \in e_2^1(j,\ell)} |\varphi(x_i)|^{1/p} \right) \cdot |D^{(\beta)} f_{j+\ell}(\psi(x_1), \dots, \psi(x_d))|,
\end{aligned}$$

see (4.6), which leads to

$$\begin{aligned}
&\left| D^{(\alpha)} \left[ \bigotimes_{i=1}^d \varphi \right](x) D^{(\beta)} [f_{j+\ell}(\psi, \dots, \psi)](x) \right| \\
&\lesssim \left( \prod_{i \in [d] \setminus e_2^2(j,\ell)} |\varphi(x_i)|^{1/p} \right) \cdot \left| \left( \prod_{i \in e_2^2(j,\ell)} \varphi^{(m)}(x_i) \right) [D^{(\beta)} f_{j+\ell}](\psi(x_1), \dots, \psi(x_d)) \right|.
\end{aligned}$$

Changing variable we obtain

$$\begin{aligned}
&\left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \left( \prod_{i \in e_2(j,\ell)} 2^{-j_i m \theta} \right) \left| D^{(\alpha)} \left[ \bigotimes_{i=1}^d \varphi \right] \cdot D^{(\beta)} [f_{j+\ell}(\psi, \dots, \psi)] \right|^\theta \right)^{1/\theta} \right\|_p \\
&\lesssim \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \left( \prod_{i \in e_2(j,\ell)} 2^{-j_i m \theta} \right) \left| \left( \prod_{i \in e_2^2(j,\ell)} \varphi^{(m)}(x_i) \right) [D^{(\beta)} f_{j+\ell}](x, \psi_{e_2^2(j,\ell)}) \right|^\theta \right)^{1/\theta} \right\|_p.
\end{aligned} \tag{4.21}$$

Here  $[D^{(\beta)} f_{j+\ell}](x, \psi_{e_2^2(j, \ell)}) = [D^{(\beta)} f_{j+\ell}](z_1, \dots, z_d)$  with  $z_i = \psi(x_i)$  if  $i \in e_2^2(j, \ell)$  otherwise  $z_i = x_i$ . Using the same scaling argument as in (4.4) and (4.8) yields

$$\begin{aligned}
& \left| \left( \prod_{i \in e_2^2(j, \ell)} \varphi^{(m)}(x_i) \right) D^{(\beta)}[f_{j+\ell}](x, \psi_{e_2^2(j, \ell)}) \right| \\
& \lesssim \left( \prod_{i \in e_2^2(j, \ell)} 2^{(j_i + \ell_i)m} \right) \sup_{y \in \mathbb{R}^d} \frac{|D^\beta f_{j+\ell}(y)|}{\prod_{i=1}^d (1 + 2^{j_i + \ell_i} |x_i - y_i|)^a} \\
& \lesssim \left( \prod_{i \in e_2^2(j, \ell)} 2^{(j_i + \ell_i)m} \right) \left( \prod_{i \in e_2^2(j, \ell)} 2^{(j_i + \ell_i)\beta_i} \right) \sup_{y \in \mathbb{R}^d} \frac{|f_{j+\ell}(y)|}{\prod_{i=1}^d (1 + 2^{j_i + \ell_i} |x_i - y_i|)^a} \\
& \lesssim \left( \prod_{i \in e_2(j, \ell)} 2^{(j_i + \ell_i)m} \right) P_{2^{j+\ell}, a} f_{j+\ell}(x)
\end{aligned} \tag{4.22}$$

with  $m \geq a$ . Now we obtain from (4.20), (4.21) and (4.22)

$$\begin{aligned}
F_\ell & \lesssim \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \left( \prod_{i \in e_2(j, \ell)} 2^{\ell_i m \theta} \right) |P_{2^{j+\ell}, a} f_{j+\ell}|^\theta \right)^{1/\theta} \right\|_p \\
& \lesssim \left\| \left[ \sum_{j \in \mathbb{N}_0^d} \left( \prod_{i \in e_2(j, \ell)} 2^{\ell_i(m-s)} \right) \left( \prod_{i \in e_0(j) \cup e_1(j, \ell)} 2^{-s \ell_i} \right) 2^{|j+\ell|_1 s \theta} |P_{2^{j+\ell}, a} f_{j+\ell}|^\theta \right]^{1/\theta} \right\|_p \\
& \lesssim \left\| \left[ \sum_{j \in \mathbb{N}_0^d} \left( \prod_{i: \ell_i < 0} 2^{\ell_i(m-s)} \right) \left( \prod_{i: \ell_i \geq 0} 2^{-s \ell_i} \right) 2^{|j+\ell|_1 s \theta} |P_{2^{j+\ell}, a} f_{j+\ell}|^\theta \right]^{1/\theta} \right\|_p
\end{aligned}$$

with  $a > \max\{1/p, 1/\theta\}$ . This is because  $i \in e_0(j)$  implies  $\ell_i \geq 0$ . Then Theorem 2.4 yields

$$F_\ell \lesssim \left( \prod_{i: \ell_i < 0} 2^{\ell_i(m-s)} \right) \left( \prod_{i: \ell_i \geq 0} 2^{-s \ell_i} \right) \|f\|_{\mathbf{F}_{p, \theta}^s}.$$

*Step 3.* In a view of (4.16) we conclude that

$$\left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \mathcal{R}_m^{e(j)}(T_d^\psi f, 2^{-j}, \cdot)^\theta \right)^{1/\theta} \right\|_p \lesssim \|f\|_{\mathbf{F}_{p, \theta}^s}.$$

The proof is complete.  $\square$

**Proof of Theorem 4.1.** *Step 1.* Let  $\{\eta_j\}_{j \in \mathbb{N}_0^d}$  and  $\{f_j\}_{j \in \mathbb{Z}^d}$  be defined as in the proof of Theorem 4.2. Analogous to (4.16) we have

$$\begin{aligned}
\|T_d^\psi f\|_{\mathbf{B}_{p, \theta}^s}^{(m)} & \leq \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \left( \sum_{\ell \in \mathbb{Z}^d} \|\mathcal{R}_m^{e(j)}(T_d^\psi f, 2^{-j}, \cdot)\|_p \right)^\theta \right)^{1/\theta} \\
& \leq \left( \sum_{\ell \in \mathbb{Z}^d} \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \|\mathcal{R}_m^{e(j)}(T_d^\psi f_{j+\ell}, 2^{-j}, \cdot)\|_p^\theta \right)^{1/\theta}.
\end{aligned}$$

Here we assume  $\theta < 1$ . If  $\theta \geq 1$  we use triangle inequality.

*Step 2.* The case  $p > 1$ . We put  $m = \lfloor s \rfloor + 1$ . The proof is similar in  $\mathbf{F}$ -spaces but less technical. Here we use the classical Hardy-Littlewood maximal inequality (the scalar version

of Theorem 2.1) and the scalar version of Theorem 2.4. These inequalities do not depend on the parameter  $\theta$ . Hence the result in the  $\mathbf{B}$ -spaces can be extended to  $0 < \theta \leq \infty$ .

*Step 3.* The case  $p = 1$ . Using the convolution representation of the  $m$ th order difference in  $e_2(j, \ell)$  we obtain from (4.19)

$$\begin{aligned} \Delta_h^{m, e(j)}(T_d^\psi f_{j+\ell}, x) &= \sum_k \Delta_h^{m, e_2(j, \ell)}(T_d^\psi f_{j+\ell}, x + (k \cdot h)) \\ &= \sum_k \int_{\mathbb{R}^{|e_2(j, \ell)|}} D^{(m_{e_2(j, \ell)})}(T_d^\psi f_{j+\ell})(x + y + (k \cdot h)) \prod_{i \in e_2(j, \ell)} h_i^{m-1} M_m(h_i^{-1} y_i) dy_i. \end{aligned}$$

Here  $y = (y_1, \dots, y_d)$  with  $y_i = 0$  if  $i \notin e_2(j, \ell)$ . We used the same notation in the proof of Theorem 4.2, see (4.17) and (4.18). Inserting this into  $\|\mathcal{R}_m^{e(j)}(T_d^\psi f, 2^{-j}, \cdot)\|_1$  and then changing the order of integration with the fact that

$$\prod_{i \in e_2(j, \ell)} \int_{\mathbb{R}} h_i^{-1} M_m(h_i^{-1} y_i) dy_i = 1$$

we obtain

$$\begin{aligned} \|\mathcal{R}_m^{e(j)}(T_d^\psi f, 2^{-j}, \cdot)\|_1 &\lesssim \left( \prod_{i \in e_2(j, \ell)} 2^{-j_i m} \right) \int_{\Gamma_j^{e(j)}} \|D^{(m_{e_2(j, \ell)})}(T_d^\psi f_{j+\ell})(\cdot)\|_1 \prod_{i \in e(j)} 2^j dh_i \\ &\lesssim \left( \prod_{i \in e_2(j, \ell)} 2^{-j_i m} \right) \|D^{(m_{e_2(j, \ell)})}(T_d^\psi f_{j+\ell})(\cdot)\|_1. \end{aligned}$$

Next step is carried out as  $\mathbf{F}$ -spaces. Note that in this case we choose  $m = \lfloor s \rfloor + 2$ , since there must exist  $a$  such that  $1 < a \leq m$  for the inequality (4.22) to hold. The proof is complete.  $\square$

**Remark 4.7.** The last step of the proof for  $p = 1$  shows that, based on our method, we have to guarantee that  $\max\{1, s\} < m$ . That's why we need the more restrictive condition  $k > m + 1 = \lfloor s \rfloor + 3$  in Theorem 1.1. Under the additional assumption  $s \geq 1$  we can relax this condition to  $k > \lfloor s \rfloor + 2$  as in the case  $p > 1$ .

Another possibility to relax this condition is to use Nikol'skij inequality at the appropriate point, as it was already done in [8].

## 5 Pointwise multiplication

We will first comment on the result for  $\mathbf{F}$ -spaces in Theorem 1.4. Again, we make use of the useful characterization by differences in Theorem 3.7. The technique used in the proof might not be “optimal” in the sense of Remarks 5.2, 5.3 below. However, it is quite transparent and works well for spaces with dominating mixed smoothness  $\mathbf{A}_{p, \theta}^s$  where almost nothing is known in this direction. Recall the multiplication operator  $\tilde{T}_d^\psi$  is defined by

$$\tilde{T}_d^\psi: f(x) \mapsto \psi(x) \cdot f(x), \quad f \in L_1(\mathbb{R}^d), \quad x \in \mathbb{R}^d,$$

for some sufficiently smooth and compactly supported function  $\psi$ . Similar to Lemma 4.4 we will proof a one-dimensional version of Theorem 1.4 first.



**Lemma 5.1.** Let  $0 < p < \infty$ ,  $0 < \theta \leq \infty$  and  $s > \sigma_{p,\theta}$ . Let further  $\psi \in C_0^k(\Omega)$  with  $k \geq \lfloor s \rfloor + 1$  for some compact set  $\Omega \subset \mathbb{R}$ . Then

$$\|\tilde{T}_1^\psi f\|_{F_{p,\theta}^s} \lesssim \|f\|_{F_{p,\theta}^s(\mathbb{T})}, \quad f \in F_{p,\theta}^s(\mathbb{T}). \quad (5.1)$$

**Remark 5.2** (Small values of  $s$ ).

(i) We believe that the conditions  $s > \sigma_{p,\theta}$  and  $s > \sigma_p$  in Lemma 5.1 and Theorems 1.3, 1.4 are only technical and caused by our proof technique. On the one hand  $s > \sigma_{p,\theta}$  (and  $s > \sigma_p$  in the  $B$ -case) is necessary for the characterization of Triebel-Lizorkin spaces by means of differences, see Theorems 3.6, 3.7 and [4], and on the other hand our proof below requires this condition. One may relax (or even remove) this condition by using the characterization of Triebel-Lizorkin spaces via local means, see Proposition 3.8 above, [34, Thm. 4.2.2] and [42, Thm. 1.25]. Note that removing the condition  $s > \sigma_{p,\theta}$  will extend the range of parameters for which (1.13) holds. In fact, if  $1/\theta - 1 > 1/p$  then the case  $1/p \leq s < 1/\theta - 1$  for the  $\mathbf{F}$ -scale in (1.13) is not covered by Theorem 1.4, although it is reasonable since  $\mathbf{F}_{p,\theta}^s \hookrightarrow C(\mathbb{R}^d)$ . In other words, we would obtain

$$\text{Int}_n(\dot{\mathbf{A}}_{p,\theta}^s) \asymp \text{Int}_n(\mathbf{A}_{p,\theta}^s(\mathbb{T}^d)), \quad n \in \mathbb{N},$$

holds true *whenever*  $\mathbf{A}_{p,\theta}^s$  is embedded into  $C(\mathbb{R}^d)$  which represents a certain minimal condition such that (1.1) makes sense.

(ii) The condition  $s > \sigma_p$  ensures that the spaces  $\mathbf{A}_{p,\theta}^s$  consist of regular distributions represented by a locally integrable function, see Lemma 3.5, (i) above. Therefore, the pointwise multiplication is well-defined. However, inspecting the proofs below we only need to give sense to the pointwise multiplication for smooth building blocks, see (4.15) above, such that we may define

$$\tilde{T}_d^\psi f := \sum_{j \in \mathbb{N}_0^d} \psi(x) f_j(x), \quad (5.2)$$

where convergence is considered in  $\mathcal{S}'(\mathbb{R}^d)$ . From that point of view one may even drop the condition  $s > \sigma_p$  at the price of  $k = k(p, s)$  in both,  $B$  and  $F$ -spaces. We leave the details to the interested reader.

**Remark 5.3** (Pointwise multiplier spaces and optimality).

(i) We ask for a sharp description of the space of pointwise multipliers for spaces  $M(\mathbf{A}_{p,\theta}^s(\mathbb{R}^d))$ . This space might be much larger than  $C_0^k(\Omega)$ , which we considered in Theorem 1.4. Questions of this type have some history, see [23, Chapt. 4] and the references therein. In that sense, the result in Lemma 5.1 is far from being optimal. In fact, combining the technique from [23, Thm. 4.9.1] with [35, Thms. 1.26, 1.29] one can prove the following rather sharp result: If  $A_{p,\theta}^s(\mathbb{R}^d)$  (which is the isotropic version of (3.1)) embeds into  $C(\mathbb{R}^d)$  then  $\psi$  is a pointwise multiplier in the sense above if  $\psi \in \dot{A}_{p,\theta}^s(\Omega)$ . What concerns the case  $s < n/p$  we refer to [25].

(ii) The technique used in [23, Chapt. 4] is not yet developed for spaces with dominating mixed smoothness  $\mathbf{A}_{p,\theta}^s$ . This seems to be a difficult task and is not straight-forward. In fact, it is for instance not clear whether all  $\psi \in \dot{A}_{p,\theta}^s(\Omega)$  are pointwise multipliers in  $\mathbf{A}_{p,\theta}^s \hookrightarrow C(\mathbb{R}^d)$ .

## 5.1 Proof of Lemma 5.1

For the proof we need the following variant of [39, Lemma 3.3.1].

**Lemma 5.4.** Let  $a > 0$ ,  $b \geq 1$ ,  $m \in \mathbb{N}$ ,  $h \in \mathbb{R} \setminus \{0\}$ ,  $\psi \in C_0^k(\mathbb{R})$  with  $k \geq m$  and  $f \in \mathcal{S}'(\mathbb{R})$  such that  $\text{supp}(\mathcal{F}f) \subset [-b, b]$ . Then it holds for any  $t \in \mathbb{R}$

$$|\Delta_h^m(\psi f, t)| \leq C_{m,a,\psi} \max\{1, |bh|^a\} \min\{1, |bh|^m\} P_{b,a}f(t). \quad (5.3)$$

*Proof.* Assume first that  $b = 1$ . By the mean value theorem from calculus we get

$$\begin{aligned} |\Delta_h^m(\psi f, t)| &\leq |h|^m \max\{1, |mh|^a\} \sup_{|\xi| \leq mh} \frac{|(\psi f)^{(m)}(t - \xi)|}{(1 + |\xi|)^a} \\ &\leq c_{m,a} |h|^m \max\{1, |h|^a\} \sup_{\xi \in \mathbb{R}} \frac{|\sum_{j=0}^m \binom{m}{j} \psi^{(j)}(t - \xi) f^{(m-j)}(t - \xi)|}{(1 + |\xi|)^a}, \end{aligned}$$

where we used Leibniz' rule in the second estimate. Let us define

$$c_{m,\psi} := 2^m \max_{j=0,\dots,m} \|\psi^{(j)}\|_\infty \quad (5.4)$$

and continue with triangle inequality and (2.3) to obtain

$$|\Delta_h^m(\psi f)| \leq c_{m,a} \cdot c_{m,\psi} \cdot |h|^m \max\{1, |h|^a\} \cdot \sup_{\xi \in \mathbb{R}} \frac{|f(t - \xi)|}{(1 + |\xi|)^a}, \quad x \in \mathbb{R}. \quad (5.5)$$

What follows is a simple homogeneity argument to deal with  $b > 1$  and  $\text{supp}(\mathcal{F}f) \subset [-b, b]$ . Putting  $\psi_b := \psi(\cdot/b)$  and  $f_b := f(\cdot/b)$  we get  $\Delta_h^m(\psi f) = \Delta_{bh}^m(\psi_b f_b)(bt)$ . Further, we have  $c_{m,\psi_b} \lesssim c_{m,\psi}$  in (5.4). Due to  $\text{supp}(\mathcal{F}f_b) \subset [-1, 1]$  we can apply (5.5) to obtain for  $t \in \mathbb{R}$

$$\begin{aligned} |\Delta_h^m(\psi f, t)| &= |\Delta_{bh}^m(\psi_b f_b, bt)| \leq c_{m,a} \cdot c_{m,\psi} \cdot |bh|^m \max\{1, |bh|^a\} \cdot \sup_{\xi \in \mathbb{R}} \frac{|f_b(bt - \xi)|}{(1 + |\xi|)^a} \\ &= c_{m,a} \cdot c_{m,\psi} \cdot |bh|^m \max\{1, |bh|^a\} \cdot \sup_{\xi \in \mathbb{R}} \frac{|f(t - \xi)|}{(1 + |b\xi|)^a} \\ &= C_{m,a,\psi} |bh|^m \max\{1, |bh|^a\} P_{b,a}f(t). \end{aligned} \quad (5.6)$$

On the other hand, by (3.2), we observe for  $t \in \mathbb{R}$

$$\begin{aligned} |\Delta_h^m(\psi f, t)| &\leq \sum_{j=0}^m \left| \binom{m}{j} \psi(t + jh) \right| \cdot |f(t + jh)| \\ &\leq c_{m,\psi} c_{m,a} \max\{1, |bh|^a\} \sup_{|\xi| \leq mh} \frac{|f(t - \xi)|}{(1 + |b\xi|)^a} \\ &\leq C_{m,a,\psi} \max\{1, |bh|^a\} P_{b,a}f(t). \end{aligned} \quad (5.7)$$

Now, (5.6) together with (5.7) imply (5.3). □

**Proof of Lemma 5.1.** In the case  $\min\{p, \theta\} \leq 1$  we choose  $0 < \lambda < \min\{p, \theta\}$  and  $a > 1/\min\{p, \theta\}$  such that  $s - (1 - \lambda)a > 0$ . It is possible since  $s > \sigma_{p,\theta}$ . If  $\min\{p, \theta\} > 1$  we put  $\lambda = 1$ . Let  $\Omega' = \Omega + [-1, 1]$  and note that  $\text{supp}(\tilde{T}_1^\psi f) \subset \text{supp}(\psi) \subset \Omega$  by assumption. Then,

clearly with  $m \geq \lfloor s \rfloor + 1$  we have

$$\begin{aligned}
\|\tilde{T}_1^\psi f\|_{F_{p,\theta}^s}^{(m)} &\lesssim \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(\tilde{T}_1^\psi f, \cdot)| dh \right]^\theta \right)^{1/\theta} \right\|_{L_p(\Omega')} \\
&\leq \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ \sum_{\ell \in \mathbb{Z}} 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(\tilde{T}_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{1/\theta} \right\|_{L_p(\Omega')} \\
&\leq \left\| \left( \sum_{\ell < 0} \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(\tilde{T}_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{u/\theta} \right)^{1/u} \right\|_{L_p(\Omega')} \\
&\quad + \left\| \left( \sum_{\ell \geq 0} \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(\tilde{T}_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{u/\theta} \right)^{1/u} \right\|_{L_p(\Omega')}
\end{aligned} \tag{5.8}$$

where  $u = \min\{1, \theta\}$ . By (5.3) we obtain for  $\ell < 0$  that

$$2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(\tilde{T}_1^\psi f_{j+\ell}, \cdot)| dh \lesssim 2^{\ell m} P_{2^{\ell+j}, a} f_{j+\ell}(t)$$

and hence

$$\begin{aligned}
&\left\| \left( \sum_{\ell < 0} \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(\tilde{T}_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{u/\theta} \right)^{1/u} \right\|_{L_p(\Omega')} \\
&\leq \left\| \left( \sum_{\ell < 0} 2^{\ell(m-s)} \left( \sum_{j \in \mathbb{N}_0} 2^{(j+\ell)s\theta} |P_{2^{j+\ell}, a} f_{j+\ell}|^\theta \right)^{u/\theta} \right)^{1/u} \right\|_{L_p(\Omega')} \\
&\lesssim \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} |P_{2^j, a} f_j|^\theta \right)^{1/\theta} \right\|_{L_p(\mathbb{T})},
\end{aligned}$$

since the function inside the  $L_p$ -(quasi-)norm is 1-periodic. Applying the periodic version of Theorem 2.4 with  $a > \max\{1/p, 1/\theta\}$ , see [24, Thm. 3.4.1], we obtain

$$\left\| \left( \sum_{\ell < 0} \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(\tilde{T}_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{u/\theta} \right)^{1/u} \right\|_{L_p(\Omega')} \lesssim \|f\|_{F_{p,\theta}^s(\mathbb{T})}. \tag{5.9}$$

Let us deal with  $\ell \geq 0$ . We write

$$|\Delta_h^m(\psi f_{j+\ell}, t)| = |\Delta_h^m(\psi f_{j+\ell}, t)|^{1-\lambda} \cdot |\Delta_h^m(\psi f_{j+\ell}, t)|^\lambda.$$

By (5.3) we obtain

$$|\Delta_h^m(\psi f_{j+\ell}, t)|^{1-\lambda} \lesssim 2^{\ell a(1-\lambda)} |P_{2^{j+\ell}, a} f_{j+\ell}(t)|^{1-\lambda}.$$

This gives

$$2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(\psi f_{j+\ell}, t)|^\lambda dh \leq M[|f_{j+\ell}|^\lambda](t)$$

and therefore

$$\begin{aligned}
& \sum_{\ell \geq 0} \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(\tilde{T}_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{u/\theta} \\
& \lesssim \sum_{\ell \geq 0} \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} 2^{\ell a(1-\lambda)\theta} \left[ |P_{2^{j+\ell}, a} f_{j+\ell}(t)|^{1-\lambda} M[|f_{j+\ell}|^\lambda](t) \right]^\theta \right)^{u/\theta} \\
& \lesssim \sum_{\ell \geq 0} 2^{[a(1-\lambda)-s]\ell u} \left( \sum_{j \in \mathbb{N}_0} 2^{(j+\ell)s\theta} \left[ |P_{2^{j+\ell}, a} f_{j+\ell}(t)|^{1-\lambda} M[|f_{j+\ell}|^\lambda](t) \right]^\theta \right)^{u/\theta} \\
& \lesssim \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ |P_{2^j, a} f_j(t)|^{1-\lambda} M[|f_j|^\lambda](t) \right]^\theta \right)^{u/\theta}.
\end{aligned}$$

Plugging this in (5.8) and using Hölder's inequality twice with  $1/\lambda$  and  $1/(1-\lambda)$  gives

$$\begin{aligned}
& \left\| \left( \sum_{\ell \geq 0} \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(\tilde{T}_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{u/\theta} \right)^{1/u} \right\|_{L_p(\Omega')} \\
& \lesssim \left\| \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} |P_{2^j, a} f_j|^\theta \right)^{1/\theta} \right\|_{L_p(\mathbb{T})}^{1-\lambda} \cdot \left\| \left( \sum_{j \in \mathbb{N}_0} [2^{js\lambda} M(|P_{2^j, a} f_j|^\lambda)]^\theta \right)^{\frac{\lambda}{\theta}} \right\|_{L_{\frac{p}{\lambda}}(\mathbb{T})}.
\end{aligned}$$

Applying the periodic versions of Theorems 2.4 and 2.1 (note that  $\lambda < \min\{p, \theta\}$ ), see [24, Prop. 3.2.4, Thm. 3.4.1], we get

$$\left\| \left( \sum_{\ell \geq 0} \left( \sum_{j \in \mathbb{N}_0} 2^{js\theta} \left[ 2^j \int_{-2^{-j}}^{2^{-j}} |\Delta_h^m(\tilde{T}_1^\psi f_{j+\ell}, \cdot)| dh \right]^\theta \right)^{u/\theta} \right)^{1/u} \right\|_{L_p(\Omega')} \lesssim \|f\|_{F_{p, \theta}^s(\mathbb{T})}. \quad (5.10)$$

From (5.9) and (5.10) we obtain the desired estimate. The proof is complete.  $\square$

## 5.2 Proof of Theorems 1.3 and 1.4

In order to prove Theorems 1.3 and 1.4 we need a multivariate version of Lemma 5.4.

**Lemma 5.5.** Let  $b = (b_1, \dots, b_d) > 0$ ,  $a > 0$ ,  $e \subset [d]$ ,  $m \in \mathbb{N}$ ,  $\psi \in C_0^k(\mathbb{R}^d)$  with  $k \geq m$  and  $h = (h_1, \dots, h_d) \in \mathbb{R}^d$ . Let further  $f \in \mathcal{S}'(\mathbb{R}^d)$  with  $\text{supp}(\mathcal{F}f) \subset Q_b$ , where

$$Q_b := [-b_1, b_1] \times \dots \times [-b_d, b_d].$$

Then there exists a constant  $C > 0$  (independent of  $f$ ,  $b$  and  $h$ ) such that

$$|\Delta_h^{m, e}(\psi f, x)| \leq C_{m, a, \psi} \prod_{i \in e} \max\{1, |b_i h_i|^a\} \min\{1, |b_i h_i|^m\} \cdot P_{b, a}^e f(x)$$

holds for all  $x \in \mathbb{R}^d$ . Here

$$P_{b, a}^e f(x) := \sup_{y_i \in \mathbb{R}, i \in e} \frac{|f(y)|}{\prod_{i \in e} (1 + |b_i(x_i - y_i)|)^a}$$

denotes the counterpart of (2.2) for  $e \subset [d]$ .

*Proof.* The proof goes along the same lines of the proof of Lemma 5.4, see also [39, Lem. 3.3.2].  $\square$

Now we are in position to prove Theorems 1.3 and 1.4.

**Proof of Theorem 1.4.** We will show that with the assumption in Theorem 1.4 we have

$$\left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{s|j|1^\theta} \mathcal{R}_m^{e(j)}(\tilde{T}_d^\psi f, 2^{-j}, \cdot)^\theta \right)^{1/\theta} \right\|_{L_p(\mathbb{R}^d)} \lesssim \|f\|_{\mathbf{F}_{p,\theta}^s(\mathbb{T}^d)} \quad , \quad f \in \mathbf{F}_{p,\theta}^s(\mathbb{T}^d), \quad (5.11)$$

where  $\mathcal{R}_m^{e(j)}(\psi f, 2^{-j}, \cdot)$  is defined in (3.3). Let us first prove the case  $\min\{p, \theta\} \leq 1$ . We choose  $0 < \lambda < \min\{p, \theta\}$  and  $a > 1/\min\{p, \theta\}$  as in the proof of Lemma 5.1. Using the decomposition of unity,  $\text{supp}(\psi) \subset \Omega$  and  $\Omega' = \Omega + [-1, 1]^d$  we get

$$\begin{aligned} & \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|1^{s\theta}} \mathcal{R}_m^{e(j)}(\tilde{T}_d^\psi f, 2^{-j}, \cdot)^\theta \right)^{1/\theta} \right\|_p \\ & \leq \left\| \sum_{\ell \in \mathbb{Z}^d} \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|1^{s\theta}} \mathcal{R}_m^{e(j)}(\psi f_{j+\ell}, 2^{-j}, \cdot)^\theta \right)^{1/\theta} \right\|_{L_p(\Omega')}, \end{aligned}$$

see (4.16). For simplicity we assume that  $\theta \geq 1$ . Note, that  $p < 1$  is still possible here. The modifications in case  $\theta < 1$  are straight-forward, see the proof of Lemma 5.1. From (3.3) we have for  $x \in \mathbb{R}^d$

$$2^{|j|1^s} \mathcal{R}_m^{e(j)}(\psi f_{j+\ell}, 2^{-j}, x) \lesssim 2^{|j|1^s} \int_{\Gamma_j^{e(j)}} |\Delta_h^{m,e(j)}(\psi f_{j+\ell}, x)| \prod_{i \in e(j)} 2^{j_i} dh_i,$$

where  $\Gamma_j^{e(j)}$ ,  $e_1(j, \ell)$  and  $e_2(j, \ell)$  are defined above in (4.17) and (4.18). Lemma 5.5 yields

$$\begin{aligned} \Delta_h^{m,e(j)}(\psi f_{j+\ell}, x) &= \Delta_h^{m,e_1(j,\ell)} \circ \Delta_h^{m,e_2(j,\ell)}(\psi f_{j+\ell}, x) \\ &\lesssim \left( \prod_{i \in e_2(j,\ell)} |2^{j_i+\ell_i} h_i|^m \right) \Delta_h^{m,e_1(j,\ell)}(P_{2^{j+\ell},a}^{e_2(j,\ell)} f_{j+\ell}, x). \end{aligned} \quad (5.12)$$

Applying Lemma 5.5 again, see also [40, Lemma 1.12], but this time with the components in  $e_1(j, \ell)$ , we get

$$\begin{aligned} & |\Delta_h^{m,e_1(j,\ell)}(P_{2^{j+\ell},a}^{e_2(j,\ell)} f_{j+\ell}, x)| \\ &= |\Delta_h^{m,e_1(j,\ell)}(P_{2^{j+\ell},a}^{e_2(j,\ell)} f_{j+\ell}, x)|^{1-\lambda} \cdot |\Delta_h^{m,e_1(j,\ell)}(P_{2^{j+\ell},a}^{e_2(j,\ell)} f_{j+\ell}, x)|^\lambda \\ &\lesssim \left[ \left( \prod_{i \in e_1(j,\ell)} |2^{j_i+\ell_i} h_i|^a \right) P_{2^{j+\ell},a}^{e(j)} f_{j+\ell}(x) \right]^{1-\lambda} \cdot |\Delta_h^{m,e_1(j,\ell)}(P_{2^{j+\ell},a}^{e_2(j,\ell)} f_{j+\ell}, x)|^\lambda. \end{aligned}$$

Inserting this into (5.12) and performing the integration yields

$$\begin{aligned} & 2^{|j|1^s} \mathcal{R}_m^{e(j)}(\psi f_{j+\ell}, 2^{-j}, x) \\ & \lesssim 2^{|j|1^s} \left( \prod_{i \in e_2(j,\ell)} 2^{\ell_i m} \right) \left( \prod_{i \in e_1(j,\ell)} 2^{\ell_i a(1-\lambda)} \right) [P_{2^{j+\ell},a}^{e(j)} f_{j+\ell}(x)]^{1-\lambda} \cdot M_{e_1(j,\ell)}(|P_{2^{j+\ell},a}^{e_2(j,\ell)} f_{j+\ell}(x)|^\lambda) \\ & \lesssim 2^{|j|1^s} \left( \prod_{i \in e_2(j,\ell)} 2^{\ell_i m} \right) \left( \prod_{i \in e_1(j,\ell)} 2^{\ell_i a(1-\lambda)} \right) [P_{2^{j+\ell},a} f_{j+\ell}(x)]^{1-\lambda} \cdot M_{[a]}(|P_{2^{j+\ell},a} f_{j+\ell}(x)|^\lambda). \end{aligned}$$

Observe that

$$\begin{aligned}
& 2^{|j|_1 s} \left( \prod_{i \in e_2(j, \ell)} 2^{\ell_i m} \right) \left( \prod_{i \in e_1(j, \ell)} 2^{\ell_i a(1-\lambda)} \right) \\
&= 2^{|j|_1 s} \left( \prod_{i \in e_2(j, \ell)} 2^{\ell_i(m-s)} \right) \left( \prod_{i \in e_1(j, \ell)} 2^{\ell_i[a(1-\lambda)-s]} \right) \left( \prod_{i \in e_0(j)} 2^{-\ell_i s} \right) \\
&\leq 2^{|j|_1 s} \left( \prod_{i \in e_2(j, \ell)} 2^{\ell_i(m-s)} \right) \left( \prod_{i \in e_1(j, \ell) \cup e_0(j)} 2^{-\ell_i \delta} \right) \\
&= 2^{|j|_1 s} \left( \prod_{i: \ell_i < 0} 2^{\ell_i(m-s)} \right) \left( \prod_{i: \ell_i \geq 0} 2^{-\ell_i \delta} \right),
\end{aligned}$$

where  $\delta = \min\{s - a(1 - \lambda), s\} > 0$ . This allows us to estimate

$$\begin{aligned}
& \sum_{\ell \in \mathbb{Z}^d} \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \mathcal{R}_m^{e(j)}(\psi f_{j+\ell}, 2^{-j}, \cdot)^\theta \right)^{1/\theta} \\
&\lesssim \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \left[ |P_{2^j, a} f_j(x)|^{1-\lambda} \cdot M_{[d]}(|P_{2^j, a} f_j(x)|^\lambda) \right]^\theta \right)^{1/\theta}.
\end{aligned}$$

Applying Hölder's inequality twice with the pair  $(1/(1-\lambda), 1/\lambda)$  and use afterwards the periodic version of Theorem 2.1, see [39, Thm. 4.1.2] and [40, Prop. 1.3], we arrive at

$$\begin{aligned}
& \left\| \sum_{\ell \in \mathbb{Z}^d} \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} \mathcal{R}_m^{e(j)}(\psi f_{j+\ell}, 2^{-j}, \cdot)^\theta \right)^{1/\theta} \right\|_{L_p(\Omega')} \\
&\lesssim \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} |P_{2^j, a} f_j|^\theta \right)^{1/\theta} \right\|_{L_p(\mathbb{T}^d)}^{1-\lambda} \cdot \left\| \left( \sum_{j \in \mathbb{N}_0^d} [2^{|j|_1 s \lambda} M_{[d]}(|P_{2^j, a} f_j|^\lambda)]^{\frac{\theta}{\lambda}} \right)^{\frac{\lambda}{\theta}} \right\|_{L_{\frac{p}{\lambda}}(\mathbb{T}^d)} \\
&\lesssim \left\| \left( \sum_{j \in \mathbb{N}_0^d} 2^{|j|_1 s \theta} |P_{2^j, a} f_j|^\theta \right)^{1/\theta} \right\|_{L_p(\mathbb{T}^d)},
\end{aligned}$$

since the function inside the  $L_p$ -(quasi-)norm is 1-periodic. The periodic version of Theorem 2.4, see [39, Thm. 4.1.3] and [40, Prop. 1.5], implies the inequality (5.11). The case  $\min\{p, \theta\} > 1$  is less technical. We choose  $\lambda = 1$ . Next we take the integral in (5.12) and then estimate above by the Hardy-Littlewood maximal function according to the directions in  $e_1(j, \ell)$ . The rest is carried out similar to the case  $\min\{p, \theta\} \leq 1$ . The proof is complete.  $\square$

**Remark 5.6.** The proof of Theorem 1.3 (**B**-case) is similar but technically less involved. As in the **F**-case we choose  $0 < \lambda < p$  and  $a > 1/p$  such that  $s - a(1 - \lambda) > 0$  if  $p \leq 1$  otherwise we put  $\lambda = 1$ . Again, one can do so because of  $s > \sigma_p$ . Note, that the weaker condition  $s > \sigma_p$  comes from the fact that the Peetre maximal function (2.2) is bounded in  $L_p(\mathbb{T}^d)$ ,  $0 < p < \infty$ , if  $a > 1/p$ , which is the scalar version of Theorem 2.4.

## Acknowledgments

The authors acknowledge the fruitful discussions with D. Bazarkhanov, W. Sickel, V.N. Temlyakov, H. Triebel, and Dinh Dũng on this topic, especially at the ICERM Semester Programme “High-Dimensional Approximation” in Providence, 2014, where this project has

been initiated, and at the conference “Approximation Methods and Function Spaces” in Hasenwinkel, 2015. The authors would particularly thank D. Bazarkhanov for pointing out reference [8] to us. Moreover, Tino Ullrich gratefully acknowledges support by the German Research Foundation (DFG) and the Emmy-Noether programme, Ul-403/1-1.

## References

- [1] T. I. Amanov, *Spaces of Differentiable Functions with Dominating Mixed Derivatives*. Nauka Kaz. SSR, Alma-Ata, 1976.
- [2] N. S. Bakhvalov, Optimal convergence bounds for quadrature processes and integration methods of Monte Carlo type for classes of functions. *Zh. Vychisl. Mat. i Mat. Fiz.*, 4(4):5–63, 1963.
- [3] V. A. Bykovskii, On the correct order of the error of optimal cubature formulas in spaces with dominant derivative, and on quadratic deviations of grids info Preprint, Computing Center Far-Eastern Scientific Center, Akad. Sci. USSR, Vladivostok, 1985.
- [4] M. Christ and A. Seeger, Necessary conditions for vector-valued operator inequalities in harmonic analysis. *Proc. London Math. Soc.* (3) 93, no. 2, 447–473, 2006.
- [5] R. A. DeVore and G. G. Lorentz, *Constructive Approximation*, Springer-Verlag, New York, 1993.
- [6] J. Dick and F. Pillichshammer, *Discrepancy Theory and Quasi-Monte Carlo Integration*, to appear in: W. W. L. Chen, A. Srivastav, G. Travaglini, Springer Verlag, 2013.
- [7] V. V. Dubinin, Cubature formulas for classes of functions with bounded mixed difference, *Matem. Sbornik*, 183 (1992), issue 7; English transl. in *Math. USSR Sbornik* 76 (1993), 283–292.
- [8] V. V. Dubinin, Cubature formulae for Besov classes, *Izvestiya Math* 61(2) (1997), 259–283.
- [9] D. Dũng and T. Ullrich, Lower bounds for the integration error for multivariate functions with mixed smoothness and optimal Fibonacci cubature for functions on the square, *Math. Nachrichten*, DOI 10.1002/mana.201400048, 2014.
- [10] C. Fefferman and E. M. Stein, Some maximal inequalities, *Amer. J. Math.* 93, 107–115, 1971.
- [11] K. K. Frolov, Upper error bounds for quadrature formulas on function classes, *Dokl. Akad. Nauk SSSR* 231, 818–821, 1976.
- [12] P. Glasserman, *Monte Carlo Methods in Financial Engineering*. Applications of mathematics : stochastic modelling and applied probability. Springer, 2004.
- [13] A. Hinrichs, Discrepancy of Hammersley points in Besov spaces of dominating mixed smoothness. *Math. Nachr.*, 283(3):478–488, 2010.
- [14] A. Hinrichs, L. Markhasin, J. Oettershagen, T. Ullrich, Optimal quasi-Monte Carlo rules on order 2 digital nets for the numerical integration of multivariate periodic functions. *Num. Math.*, to appear.
- [15] E. Hlawka, Zur angenäherten Berechnung mehrfacher Integrale. *Monatsh. Math.*, 66:140–151, 1962.
- [16] N. M. Korobov, Approximate evaluation of repeated integrals. *Dokl. Akad. Nauk SSSR*, 124:1207–1210, 1959.
- [17] D. Krieg and E. Novak, A universal algorithm for multivariate integration, *ArXiv e-prints*, 2015. arXiv:1507.06853 [math.NA].

- [18] F. Kuo, I. H. Sloan and H. Woźniakowski, Periodization strategy may fail in high dimensions, *Numerical Algorithms* 46(4):369–391, 2007.
- [19] L. Markhasin, Discrepancy and integration in function spaces with dominating mixed smoothness, *Dissertationes Math.* 494, 1–81, 2013.
- [20] S. M. Nikol’skij, *Approximation of Functions of Several Variables and Embedding Theorems*. Nauka Moskva, 1977.
- [21] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems*, Volume II: Standard Information for Functionals, European Math. Soc. Publ. House, Zürich, 2010.
- [22] D. Nuyens and R. Cools, *Higher Order QuasiMonte Carlo Methods: A Comparison*, AIP Conf. Proc. 1281, 553–557, 2010.
- [23] T. Runst and W. Sickel, *Sobolev Spaces of Fractional Order, Nemytskij Operators, and Nonlinear Partial Differential Equations*, volume 3 of *de Gruyter Series in Nonlinear Analysis and Applications*. Walter de Gruyter & Co., Berlin, 1996.
- [24] H.-J. Schmeisser and H. Triebel, *Topics in Fourier Analysis and Function Spaces*, Geest & Portig, Leipzig, 1987 and Wiley, Chichester, 1987.
- [25] W. Sickel, On pointwise multipliers for  $F_{p,q}^s$ . The case  $\sigma_{p,q} < s < n/p$ , *Annali di Matematica pura ed applicata* 174, 209–250, 1999.
- [26] M. M. Skriganov, Constructions of uniform distributions in terms of geometry of numbers, *Algebra i Analiz* 6, 200–230, 1994.
- [27] B. Stöckert, Ungleichungen vom Plancherel-Polya-Nikolskij-Typ in gewichteten  $L_p^\Omega$ -Räumen mit gemischten Normen, *Math. Nachr.* 86, 19–32, 1978.
- [28] V. N. Temlyakov, On reconstruction of multivariate periodic functions based on their values at the knots of number-theoretical nets. *Anal. Math.*, 12:287–305, 1986.
- [29] V. N. Temlyakov, On a way of obtaining lower estimates for the errors of quadrature formulas. *Mat. Sb.*, 181(10):1403–1413, 1990.
- [30] V. N. Temlyakov, Error estimates for Fibonacci quadrature formulas for classes of functions with a bounded mixed derivative. *Trudy Mat. Inst. Steklov.*, 200:327–335, 1991.
- [31] V. N. Temlyakov, *Approximation of Periodic Functions*. Computational Mathematics and Analysis Series, Nova Science Publishers, Inc., Commack, NY, 1993.
- [32] V. N. Temlyakov, Cubature formulas, discrepancy, and nonlinear approximation, *J. Complexity* 19, 352–391, 2003.
- [33] H. Triebel, *Theory of Function Spaces*. Birkhäuser, Basel, 1983.
- [34] H. Triebel, *Theory of Function Spaces II*. Birkhäuser, Basel, 1992.
- [35] H. Triebel, *Function spaces and wavelets on domains*. EMS Publ. House, Zürich, 2008.
- [36] H. Triebel, *Bases in Function Spaces, Sampling, Discrepancy, Numerical Integration*. EMS Tracts in Mathematics, European Mathematical Society (EMS), Zürich, 2010.
- [37] M. Ullrich, On ”Upper error bounds for quadrature formulas on function classes” by K. K. Frolov, *ArXiv e-prints* 1404.5457, 2014.
- [38] M. Ullrich and T. Ullrich, The role of Frolov’s cubature formula for functions with bounded mixed derivative, Preprint, 2015.
- [39] T. Ullrich, Function spaces with dominating mixed smoothness; characterization by differences. *Jenaer Schriften zur Mathematik und Informatik*, Math/Inf/05/06, 2006.
- [40] T. Ullrich, Smolyaks algorithm, sparse grid approximation and periodic function spaces with dominating mixed smoothness, Thesis, Friedrich-Schiller-Universität Jena, Jena, 2007.



- [41] T. Ullrich, Local mean characterization of Besov-Triebel-Lizorkin type spaces with dominating mixed smoothness on rectangular domains, Preprint, Uni Bonn, 2008.
- [42] J. Vybírál, Function spaces with dominating mixed smoothness, *Dissertationes Math.* 436, 73 pp., 2006.
- [43] H. Yserentant, *Regularity and Approximability of Electronic Wave Functions*. Lecture Notes in Mathematics, Springer (2010).